DOCUMENT RESUME

ED 289 354 FL 017 068

AUTHOR Mackey, William F.; And Others

TITLE Polychronometric Techniques in Behaviour Analysis: Language Teaching. Publication B-163.

INSTITUTION Laval Univ., Quebec (Quebec). International Center for Research on Bilingualism.

PUB DATE 87

NOTE 145p.

PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC06 Plus Postage.

DESCRIPTORS Bilingualism; *Classroom Communication; Classroom Techniques; *Efficiency; Elementary Secondary Education; English (Second Language); Foreign Countries; French; Interaction; *Language Teachers; *Language Usage; Measurement Techniques; Second Language Instruction; Student Behavior; *Teacher Behavior; Teacher Characteristics; *Time Factors (Learning); Verbal Tests

IDENTIFIERS *Error Correction (Language)

ABSTRACT

A study examined the relationship of time, language teachers' language use, and the qualities of interaction in the language classroom. An introductory section discusses the interest in these issues and the use and development of polychronometric techniques for behavior analysis. Three studies using these techniques to analyze language teacher behaviors follow. One examined the extent to which the components of the language lesson are significant in enhancing learning. The study compared components of five lesson types presented to 11-year-old students of English as a second language. The second compared the five lesson types in terms of the relative duration and frequency of didactic and language behaviors. The third study looked at the relationship between learner language use, teacher language behavior, error correction, and the level of language learning among third- and ninth-grade students. A concluding chapter discusses the studies' findings as they relate to measuring language teaching efficiency. An epilogue addresses the persistent attitude that teaching a foreign language may really be a waste of time. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

centre international de recherche sur le bilinguisme

publication
B-163

ED289354

POLYCHRONOMETRIC TECHNIQUES IN BEHAVIOUR ANALYSIS:
Language Teaching

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

A. Prujiner

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

William F. Mackey
Glen Loveless
Perlette Louisy
Veronica Heechung

1987

CIRB
ICRB

FL017068

William F. MACKEY
Glen LOVELESS
Perlette LOUISY
Veronica HEECHUNG

POLYCHRONOMETRIC TECHNIQUES IN BEHAVIOUR ANALYSIS:

Language Teaching

Publication B-163

1987
Centre international de recherche sur le bilinguisme
International Center for Research on Bilingualism
Québec

Le Centre international de recherche sur le bilinguisme est un organisme de recherche universitaire qui reçoit une contribution du Secrétariat d'État du Canada pour son programme de publication.

The International Center for Research on Bilingualism is a university research institution which receives a contribution from the Secretary of State of Canada for its publication programme.

Ⓒ **1987 CENTRE INTERNATIONAL DE RECHERCHE SUR LE BILINGUISME**
Tous droits réservés.
Imprimé au Canada
Dépôt légal (Québec) 3^{e} trimestre 1987
ISBN 2-89219-178-5

PREFACE

Interest in the study of time variables in behaviour goes back, in our case, to the early fifties when I was attempting to quantify elements in the verbal behaviour of bilinguals. It became evident that in fact all human behaviour comprises time variables, some of which are crucial. This we hope to demonstrate in future studies.

It was only in the early sixties, however, after our *Language Teaching Analysis* had gone to press (Mackey, 1965) that the need to quantify language teaching behaviour became evident. It was needed as a sequel to our quantification of language teaching methods and materials whose application was the end-game, the bottom line, as it were, of the entire language teaching operation on which so much had been expended.

In order to be able to quantify the time variables in such a complex behaviour as language teaching it was necessary to develop the appropriate measuring instruments. This task took an initial three years of technical work in collaboration with the late, and much regretted Lucien Le Bourhis, our ingenious lab technician whose important contribution to this phase is gratefully acknowledged.

With this appropriate instrumentation, we were able to develop in the seventies and early eighties a number of effective polychronometric techniques of behaviour analysis. We tested these with the help of our students some of whom went on to earn masters and doctorate degrees for their work. This work comprises the middle chapters of this book.

First there was the work of Perlette LOUISY, who tested the techniques against others already in use (Louisy, 1975). Then came the application by Veronica HEECHUNG to the relationship between teaching and learning (Heechung, 1981). More recently, Glenn LOVELESS, in his doctorate thesis, used these techniques to measure the effectiveness of different didactic behaviours on the learning of a second language (Loveless, 1983). We hope that this is only a beginning.

Although both English and French as second languages are represented in these studies, it is obvious that these techniques of analysis are applicable to the teaching of any language.

What also had made this book possible is the contribution of a number of people and agencies. In the name of my three collaborators I should now like to thank each of them - particularly the agencies that granted the scholarships and leaves which provided the hundreds of hours of time necessary for the painstaking work of testing the techniques. These agencies include the Canadian Commonwealth Scholarships and Fellowships Committee, the Pentacostal Assemblies Board of Education in Newfoundland, the Special Services of the Government of Newfoundland and the Quebec Government.

The dozens of teachers in Newfoundland and in Quebec who allowed their lessons to be taped and filmed deserve a special thanks; without their cooperation, none of this could have been done.

Many of the ideas and applications I owe to my wife, Dr. Ilonka SCHMIDT-MACKEY, who was the first to use these techniques in her innovative courses in language teaching techniques and in the in-service training of language teachers.

Finally, I owe a special thanks to Marguerite MARCHAND, administrative officer of the ICRB, for her many hours of work, far exceeding the calls of duty, which she devoted to the preparation of the manuscript. Finally I should like to extend my warm thanks to our head of publications, Richard VIGNEAULT, for his patience, understanding and skillful editing of the text.

For mistakes of any kind that may have crept into the published version I must take full responsibility.

William F. Mackey, FRSC
ICRB, Research Professor

TABLE OF CONTENTS

INTRODUCTION:

Time in Human Behaviour 1

CHAPTER 1:

Measuring Time Variables in Behaviour: Problems and Solutions 6

CHAPTER 2:

Polychronometric Profiles of Teaching Behaviour. 13

CHAPTER 3:

Learning and Teaching Behaviour: Comparing Lesson Types 39

CHAPTER 4:

Learning and Teaching Behaviour: Correlating Lesson Components. 69

CHAPTER 5:

Do Some Types of Teaching Behaviour Yield More Learning?
Evidence from Polychronometric Experiments 101

CONCLUSION:

Measuring Language Teaching Efficiency 119

EPILOGUE:

Is Language Teaching a Waste of Time? 133

REFERENCES 137

CODES AND SIGLA 141

LIST OF TABLES

TABLE 1:
Categories Polychronometric Technique of Foreign Language Teaching Analysis 14

TABLE 2:
Teacher-Student Behaviour in Lesson A 19

TABLE 3:
Language Behaviour in Lesson A. 22

TABLE 4:
Activities: Types and Occurrences: Lesson A 23

TABLE 5:
Teacher-Student Behaviour in Lesson B 23

TABLE 6:
Language Behaviour in Lesson B. 27

TABLE 7:
Activities: Types and Occurrences: Lesson B 28

TABLE 8:
Teacher-Student Behaviour in Lesson C 29

TABLE 9:
Language Behaviour in Lesson C. 31

TABLE 10:
Activities - Types and Occurrences: Lesson C 32

TABLE 11:
Polychronometric Indices for Lessons A, B and C 33

TABLE 12:
Lesson A: Activity and Variety Indicators 41

TABLE 13:
Lesson B: Activity and Variety Indicators 42

TABLE 14:
Lesson C: Activity and Variety Indicators 43

TABLE 15:
Lesson D: Activity and Variety Indicators 44

TABLE 16:
Lesson E: Activity and Variety Indicators 45

TABLE 17:
The Relative Duration and Frequency of Components in Five Lessons 70

TABLE 18:
Sum of Duration and Frequencies of Behaviours in Five Lessons. 74

TABLE 19:
Frequencies of Language Use in Five Lessons 74

TABLE 20:
Intensity Indicators for Five Lessons 79

TABLE 21:
Correlation Indices: Behaviour Indicators and Oral Tests 83

TABLE 22:
Correlation Indices: General Behaviour Indicators and the Oral Tests 84

TABLE 23:
Correlation Indices: Activity Indicators and Oral Tests 86

TABLE 24:
Correlation Indices: Activity Indicators and Oral Tests 89

TABLE 25:
Correlation Indices: of Language Use and Oral Test 91

TABLE 26:
Summary Table of Correlations 93

TABLE 27:
Summary Table of Correlations 94

TABLE 28:
Total Distribution of General Didactic Behaviours in Lessons A1 and A2 102

TABLE 29:
Total Distribution of Presentation Behaviours in Lessons A1 and A2 103

TABLE 30:
Total Distribution of Instructor Language Use Behaviours in Lessons A1 and A2. . . . 104

TABLE 31:
Total Distribution of Learner Actions and Speech Behaviours in Lessons A1 and A2 . . 104

TABLE 32:
Total Distribution of Language Performance in Lessons A1 and A2 105

TABLE 33:
Distribution of Certain General Didactic Behaviours in Lessons F3 and F4. 106

TABLE 34:
Distribution of Presentation Behaviours PA and PT during Lessons F3 and F4 107

TABLE 35:
Total Distribution of Learner RT and TR Behaviours in Lessons F3 and F4 108

TABLE 36:
Total Distribution of PX, PT and OM Behaviours in Lessons V5 and V6. 110

TABLE 37:
Total Distribution of PR, RM and CO Behaviours During Lessons V5 and V6 110

TABLE 38:
Comparison of Levels of Significance of Effects of Variety on Learning 112

TABLE 39:
Total Distribution of Action and Speech Behaviours in School DE-37 114

TABLE 40:
Total Distribution of KS, LS and OL Behaviours Among Ninth Grade Learners 116

TABLE 41:
Distribution of Independently Generated Sentences in a Lesson 130

LIST OF FIGURES

FIGURE 1:
The Polychronometer. 12

FIGURE 2 a AND 2 b:
Frequency and Relative Duration of Teaching Behaviour in Lesson A 21

FIGURE 3:
Frequency and Relative Duration of Learning Behaviour in Lesson A 21

FIGURE 4:
Duration Profile of Lesson A 22

FIGURE 5:
Frequency Profile of Lesson A. 22

FIGURE 6:
Frequency and Relative Duration of Teaching Behaviour in Lesson B 24

FIGURE 7:
Frequency and Relative Duration of Learning Behaviour in Lesson B 25

FIGURE 8:
Duration Profile of Lesson B. 27

FIGURE 9:
Frequency Profile of Lesson B 27

FIGURE 10:
Frequency and Relative Duration of Teaching Behaviour in Lesson C 29

FIGURE 11:
Group Activity in Lesson C 30

FIGURE 12:
Individual Student Learning Behaviour in Lesson C 30

FIGURE 13:
Duration Profile of Lesson C. 33

FIGURE 14:
Frequency Profile of Lesson C. 33

FIGURE 15:
Distribution of Teaching Behaviours in Lessons A, B and C: Duration 34

FIGURE 16:
Distribution of Teaching Behaviours in Lessons A, B and C: Frequency 34

FIGURE 17:
Distribution of Learning Behaviours in Lessons A, B and C: Duration 35

FIGURE 18:
Distribution of Learning Behaviours in Lessons A, B and C: Frequency 35

FIGURE 19:
Distribution Profiles of Total Behaviours. 47

FIGURE 20:
Distribution Profiles of Articulation 49

FIGURE 21:
Distribution Profiles of Presentation 52

FIGURE 22:
Distribution Profiles of Repetition 54

FIGURE 23:
Distribution Profiles of Verification 56

FIGURE 24:
Distribution Profiles of Actions 58

FIGURE 25:
Distribution Profiles of Speech thru Speech 60

FIGURE 26:
Distribution Profiles of Action and Speech. 62

FIGURE 27:
Distribution Profiles of Language Use 64

FIGURE 28:
Distribution Profiles of Language Performance 66

FIGURE 29:
Duration of Didactic and Language Behaviours in Five Classrooms. 75

FIGURE 30:
Frequency of Didactic and Language Behaviours in Five Classrooms 76

FIGURE 31:
Quantified Behavioral Sequences of an English Lesson 126

FIGURE 32:
Behaviour Distribution Profile of the Lesson. 127

Introduction:

TIME IN HUMAN BEHAVIOUR

> Before the beginning of years
> There came to the making of man
> Time . . .
>
> Swinburne, *Atlanta in Calydon*

The studies that follow are about time and events - human events. Time is admittedly the most ubiquitous of all phenomena - and the most impenetrable - being present in every process while far transcending the years of life and of nature. It has no intervals of its own, nor peaks nor valleys, nor space between units. Yet like space it is ultimately immeasurable.

Paradoxically, it is only by measuring time that we can conceive it. And our conception is continually expanding, both in range and in precision. Our knowledge of the past ranges far beyond that of our ancestors, as does the exactness with which we can measure the present, even to the imperceptibly rapid movements within the atom. All divisions of time are, in the last analysis, essentially arbitrary; some are entirely imaginary. The earliest divisions were based on the observation of nature. Recurring seasons were grouped into years. Observations of the recurring phases of the moon provided for the division of years into months whose units could be counted in days according to the rising and setting of the sun. Beyond this, for more precise measures, it was necessary to invent units of time.

The segmentation of the day into hours and of the hour into minutes are essentially divisions of convenience. There is no reason why a minute should have only sixty seconds or that an hour should be sixty minutes and not a hundred or that a day should have 24 hours, and not ten. Except that they were arbitrarily defined as such and accepted as conventions. In the final analysis, human signs are all arbitrary.

A good example of the above is the division of the month into weeks. Originally, a week was simply a convenient interval between market days - four or more days according to climate and custom. For the early Greeks it was ten days, an interval of time which was later re-instated in France in 1792 by the Revolutionary Convention as part of its universal decimal system. Other revolutions; other systems. The Soviet Revolution engendered a five-day week in 1929, increased to a six-day week in 1932 and to a seven-day week in 1940, without reference to its origin in the story of *Genesis* recounting the creation of the world in six days by the Lord, who rested on the seventh day.

The division of the day was also a matter of convenience. In most of Europe until the late Middle Ages, three divisions sufficed - morning, daytime and night - divisions known as tides (morningtide, noontide and eventide). But in the Middle East the day had long ago been divided into hours by the early Sumerians and later by the Babylonian and Egyptian priests who, observing the stars, the moon and their sun-god Re, divided the changes in the night sky into twelve phases, the positions of the sun into ten, and, between night and day, two transition periods, dawn and dusk - for a total of twenty-four. These same observations revealed that it took 365 of these days for the recurrence of the natural phenomena of the seasons and, most importantly, of the life-giving flooding of the Valley of the Nile.

For many centuries, the division of time into hours was sufficiently precise for most everyday needs. The need to break up the hour into further intervals became necessary as the effects of the Industrial Revolution quicked the pace of everyday urban activities. Industry, commerce and transportation required the making of more and more exact schedules. Several trains could leave a station within the space of an hour - but not at the same moment. Fractions of an hour permitted a more efficient use of time. As the pace of life accelerated further the minute no longer sufficed. Further divisions were needed for such things as picking winners in close horse races. As experimental sciences like ballistics and physics required more precise units, fractions of a second - tenths, hundredths and thousandths - became current.

All this precision eventually affected the everyday lives of people. In the early part of this century, workers were paid by the day ("Another day; another dollar" as the saying went). Later they were paid by the hour. Today billing is often done in fractions of the hour. Skilled workers and even professionals are expected to devote only so many minutes per operation. So that the concept of time which was first controlled by events, has evolved, through measurement, into an object which controls events - and people.

The division of time into units is one thing; the exact duration of such units is something else. The years and days of early calendars were not each of equal length. This was due largely to the fact that the year does not divide exactly into 365 days nor into 12 equal months; nor does the month divide into 30 days. There is always that extra fraction which eventually mounts up to an embarrassing cumulative error. Approaching the beginning of our era, the Roman calendar was so much behind the solar year that, after his experience in Egypt, and on the advice of its scientists, Julius Caesar added 23 days to the year 46 B.C. decreeing that it be composed of 446 days. To help avoid further inaccuracies in the future he further ordered that an extra day be added every fourth year at the end of February.

Even in Antiquity, a need was felt to make the measurement of elapsted time independent of the vagueries of nature so that all units of the same name be of equal length. It is not surprising that the Ancient Greeks, who revered Chronos, their god of time, should have fathered the chronograph, an instrument for measuring time. They did so by controlling a steady flow of water from a large container over a water wheel linked to a revolving cylinder. Others had measured time by the movement of shadow cast by the sun. Chronographs with springs and tuning forks eventually provided more precision in response to the developing needs of mankind for in navigating the globe, operating railways, managing air travel and in perfecting the work of science. Eventually, the measurement of imperceptibly minute segments of time could be done only with the aid of the atom. Today quartz crystals specially treated to emit radiation in fixed frequencies measurable in tens of thousands of cycles per second within the atoms of caesium or hydrogen can provide measurement of time accurate within less than two minutes in the millions of years since the birth of the solar system.

This leads us to the problem of the point of departure - the beginning of measurable time. Like the point of departure of many of our measures, the zero point has been based on such observable phenomena as the freezing point of water which became zero celsius. Indeed there can be no system of measure without a zero, no time measurement without a point of beginning. This is true of historical time.

The problem is one which all literate civilizations have had to face. The Ancient Greeks started counting years from the first Olympiad (circa 776 B.C.); the Romans from the alleged date of the founding of Rome (753 B.C.); the Christians from the birth of Christ. People tend to accept the system adopted by the dominant civilization of the age. But the acceptation does not come overnight. Competing systems can co-exist. For centuries, different parts of the world operated with different calendars, the Babylonian, the Egyptian, the Julian and the Gregorian, the latter, now universal, having been adopted by all of Europe except England, which finally relented when in 1752 it was eleven days behind the rest of the world. Other civilizations have maintained their traditional counting of the years. Even today we can still celebrate different new years, like the Jewish, the Chinese and the Moslem. Again, all measures of time are arbitrary. The important thing is that they be conventional - and, if possible, appropriate to the cultural content and needs of the users.

In human affairs, including scientific research, time has been both a hindrance and an embarrassement. Since much of what physical scientists observe seems to be stable - or predictable - most of the work in the human and exact sciences has been descriptive, based as it is on the physical composition of substances and the shape and size of everything in the observable universe. When applied to such unstable subjects of observation as societies, people and their individual and group behaviour such synchronic descriptions produce distorsions (Mackey 1983).

Since the use of time is a property of organisms, it goes into the patterns of their behaviour. It accounts for the observed differences between the plodding turtoise and the sprinting jackrabbit, between the frisky kitten and the aged cat, between the first-graders in school and the adults in the night classes - indeed, between differences in the personality of individuals of the same age and background.

Time is the main determinated in the adaptation of organisms to changing, and often hostile, environments, in pattern anticipation, automatic perception and such other psychological phenomena as semantic satiation. As acceleration, time is also the main component of many macro-phenomena. If one applies heat to water, for example, the molecules move at random at a faster and faster rate, until the water reaches a point, the boiling point, when the entire mass moves in rhythmic waves. The behaviour of any one molecule may be random and erratic but that of the water is regular.

Likewise with societies, both animal and human. The behaviour of individuals may be erratic each being statistically insignificant in their abberation; but the combined activity of the group as a group is oriented. This can be seen in the behaviour of insect colonies, like bees and ants, and of birds (Wilson 1975). Similarly with human societies, where each seeks his or her own welfare, while unwittingly contributing to that of the society as a whole. Such for example is macro-economic behaviour responsible for the volatility of money, regardless of supply and demand, the action of stock-markets and the rates of inter-urban communication. Much of such behaviour is quasi-biological, having to do with the welfare of kith and kin (Mackey 1985).

Since mid-century, such social activity has become both faster and more complex. Populations have become more mobile, and more diverse. The multiplicity of options - economic, social, and cultural - have modified behaviour patterns. This is evident for example, in the industries where products of great complexity have to be serviced - electronically controlled appliances, autos and other automated devices. Secondly, the variety of products and models available in developed countries poses problems of selection and judgement. Because of these numerous material, social and cultural options, people had to budget their time to a greater and greater extent - more and more to do and less time in which to do it. In this context, the mechanization of more and more services has converted personal service into a rare commodity (Albrecht & Kembe 1985). By the 1980s people with less and less time began to realize that they had to wait longer and longer to be served. In shopping centers, garages, hospitals, bars, cinemas - wherever they went they had to wait longer to get what they needed. To answer the demand, service became more and more mechanized (automatic tellers, computer-controlled checkouts) and regimented standard procedures and fixed responses like "It was nice having you on board. Thank you for flying our airline".) (Hochschild 1983).

The promotion of consumerism and mass entertainment available at all times created demands to work fewer and fewer hours for more and more pay in face of a greater need for increased productivity. All this at a time when fewer people were producing products and - paradoxically - more and more people offering services. In the 1980s more than 85% of all new jobs were in the service industries. During this period there were 83 million service workers in Canada and the Unites States, as against 28 million in manufacturing and four million in agriculture. Most of the working population was by then made up of millions of salespersons, secretaries and professionals and hundreds of thousands of nurses, hairdressers, cashiers, waitresses and teachers.

Since all these millions were paid for their time, and at ever increasing wages, those doing the paying became concerned about their investment. They became most interested in how well the employee spent his or her time when on the job. Until jobs, or their communications element, became computerized this was not always easy to do without unwelcome intrusion. But by the mid-eighties some twenty million North Americans were already using computers in their jobs. In addition to increasing efficiency the use of computers permitted some automatic monitoring of working time since such electronic supervision required only a little extra software in the central computer (Shaiken 1984). Airline reservations computers, for example, automatically measure the idle moments of the reservations clerks. A computer dossier of each employee who uses a word processor becomes an automatic measure of that person's productivity. To large companies hiring thousands of workers to do repetitive jobs, such practices were indeed tempting (Halberstam).

Most of these measurements of working efficiency however have remained one dimensional being of the same nature as those primitive time-and-motion studies. They have been used mostly for the purpose of gaining a competitive edge. They have not been used for multidimensional work analysis to determine the behaviour best suited to the personality of the worker nor to help ease the lot of the overworked. They have been mostly mono-chronometric. Contrariwise, the polychronometric techniques described and illustrated in this book are meant, not for the control of behaviour, but for its measurement, analysis and understanding. For such analysis and understanding there is a growing need. Crowd behaviour, interpersonal interaction, the hyptonic effects of television and certain spectator sports are ill understood. We first have to know what goes on, how often and how long. For example, it had been pointed out that, for a television ad to keep the viewers attention - conscious or unconscious - something must happen at least every 15 seconds, that is, visual changes in the view, the angle, the subject, the lighting, or the picture itself. Job evaluation is another field in which quantification is important if it is to fit the right person to the right work. For example, it is important to discover whether a person's tempo of interpersonal relations is the sort as to make him happier and more efficient as a teacher, an office worker, or a salesperson. And even within a single job-category the quantifications must be sensitive enough to establish a relationship between the tempo of the individual and the tempo of the job. Personnel managers have observed that one salesperson may be happier selling a lot of small items whereas the tempo of another salesperson may fit him better for selling large, expensive items where more patience and tact may be required; for success in selling depends on what is being sold by whom. The difference in characteristics needed to succeed in two different jobs often accounts for the fact that so many promotions become failures, as is often the case when a successful production worker is promoted to the job of supervisor (Chapple & Donald 1947).

The need for quantification of personal interaction has also been seen in the analysis of interview material, including psychiatric interviews. The relative amount of speech activity on the part of the interviewee can be correlated with personality traits; so can the variability of the speed of talking (Goldman-Eisler 1954).

Verbal behaviour has also been analysed and described in terms of time variables. It has been noted for example, in the average male-female conversation, that the man is responsible for as much as 96% of all interruptions whereas the woman is the author of most of the questions asked - three times more than the man (Kramarae 1981).

Interaction analysis has been used in the evaluation of teaching performance (Flanders 1970). The relation between the amount of teacher and student talk has been considered significant. So has the distribution within a lesson of direct and indirect teaching - of lecturing, criticizing or giving directions, as opposed to questioning, encouraging, making use of the students' ideas and acknowledging their feelings. The amount of talk initiated by the student as compared to the amount of talk resulting from direct reply to the teacher's questions has also been used as an index to characterize a lesson.

Quantifiable generalization and comparison of such behavioural differences are not possible, however, until some sort of standard measures are developed and applied. For the history of science is largely the history of measurement. Most developments in measurement have in the past been associated with the physical sciences. Since the turn of the century, however, advances in the social sciences have been associated with statistical methods. Most of these have been applied to the behaviour of populations, rather than to detailed observations of individuals or small groups, and their behaviour during observable events in their lives. It is the purpose of the following chapter to present a method and an instrument for the analysis and measurement of such observable events.

Although there are established ways of measuring changes in environment, there are as yet few standard measures for variables of events that take place in the environment. Environmental changes in such things as temperature, humidity, radiation, and the purity of air and water are measured daily, and these measurements are used to solve a host of practical problems. We no longer rely on our feelings to find out whether a place is warmer or damper than another; we rely on such instruments as thermometers, hydrometers, barometers, and the like. The use of such instruments permits us to record measurements of things that are continually changing, enabling us to come to a number of general and practical conclusions.

On the other hand, in talking about events we have very few possibilities of measurement. A notable exception is the stock market where the selling and buying activity of individuals is recorded with precision and rapidity. In less than a few hours we can know how many shares have changed hands in such cities as London, New York, Toronto and Tokyo, the value of these shares, and whether their price has gone up or down. But when talking about such fields as radio and television, sports, theater, ballet, and other forms of activity, we have nothing more than individual impressions and opinions, which may vary from individual, when analysing the same event. Opinions may vary as to how much action there was in one event as opposed to another, whether someone's performance was better than someone else's, and the like, because we have no objective way of analysing and measuring the variables. The same symphony, opera or play presented a number of times, each time with changes in the tempo of its components can give a series of entirely different impressions. In reacting to the merits of each we are likely to shrug our shoulders and resign ourselves to the dictum: *de gustibus non est disputandum.*

In observing people, we can get some idea of their mutual inter-relationship simply from the duration and frequency of the contact between them and the distribution of their speech and gestures. These, it would seem, should, with the aid of operant methods, lend themselves to behavioural analysis. And such analysis when quantified may well have predictive power, since the behaviour of organisms tends to follow fixed patterns.

It would seem, therefore, that elaboration and improvement of methods of quantification of time variables in observable phenomena could result in advances in the behavioural sciences. The variables, however, must first be isolated and selected.

Chapter 1

MEASURING TIME VARIABLES IN BEHAVIOUR: PROBLEMS AND SOLUTIONS

All events have variables. These variables can be isolated and subjected to measurement. They can be isolated in relation to the problems we have to solve, or the information we seek. They can be analysed as functions of such evaluative concepts as stability, stimulation and interest. These functions can be measured on a time scale, giving us the amount of time and its distribution for each variable.

1. The Problems

Before any sort of evaluative conclusions are possible, however, the event must first be accurately observed, the variables isolated and measured, the results interpreted and the functional categories established. There are therefore four related problems the solution of which depends on the field, the situation and the objectives. They are: the problem of observation, the problem of computation, the problem of interpretation and the problem of categorization.

1.1 *The Problem of Observation*

Whether the event is observed directly as it is taking place or indirectly as a recording or verbatim report may depend on the possibility of obtaining information that is accurate or abundant enough to be worth while.

Direct observation - known or unknown to the subjects being observed - was all that was possible at the period when time-and-motion studies were first launched in industry. The image of the time-and-motion observer has disturbed labor organizers. As late as the 1960s, aged trade unionists were still haunted by visions of the furtive efficiency expert with his stop watch and notebook. Yet no one seemed to the object to tallying of sports analysts and drama adjudicators.

As recording facilities became more generally avalaible, indirect observation through recorded material gained in popularity. These facilities permitted a choice between verbal recording, vocal recording, cinematography and videotaping.

Verbal records, especially the verbatim stenographic transcripts of the court clerk have been of limited use in the study of duration and frequency of events, since they have recorded only what can be written according to the usual conventions. Vocal recording, however, by means of disk and tape recorders, registers not only what is said but also how it is said, including the hesitations and silences, in real time. These have been used for the analysis of psychiatric interviews (Kasl & Mahl 1965). If, however, one is interested more in the actions than in the speech of the subjects, simple cinematographic records may be what is required. This type of recording may also be used in the analysis of interview material (Sainsbury 1954). The most complete record, of course, includes both the sound and the image. Both sound motion pictures, and especially videotape have become widely accessible to record for later analysis events in many different fields (Mackey 1968a). The superiority of videotape is that it gives the choice of either erasing the record once it has been analysed, or of conserving it as a document or a model for purposes of training and research (Mackey 1968b).

The means of observing the event - directly or indirectly - is only the beginning of the problem of description and analysis. Once we have the record, we must arrive at a means of compiling and computing the relevant data which it contains.

1.2 *The Problem of Computation*

What are the objectively measurable components of an event? The chief ones are time, sound, movement and contact. Since both sound and movement must take place in time, they can be described and measured in temporal terms as continua of frequency of occurrence and duration. All phenomena distributed in time can be described in terms of frequency distribution and relative duration; these can be correlated with related phenomena as being significant, causal or diagnostic.

In analysing configurations of movement and sound as they unroll in time, one must know what units to look for and how to count them. The units themselves will depend on the type of event, the field of knowledge and the purposes of the analysis. If one is studying interview transactions one may be interested only in what a person says (content analysis) (Fliess 1948). Or one may be interested in how he says it (behavioural analysis), and this could include the person's rate of speech, the amount of talking and the length and number of silences and interruptions.

A complete count of the time-units includes a record of their duration and frequency. The duration can be computed in arbitrary units of time - in minutes or seconds. Some researchers have used units of one minute (Melbin 1954); others have been able to limit themselves to 30 seconds (Verzeano 1949), five seconds, and even to three seconds (Moskowitz 1970). Or the duration may be computed in real time (Starkweather 1960). Both methods can be used with or without special recorders of elapsed time.

Frequency can be counted with the aid of special mechanical or electromagnetic counters (Matarazzo 1956), or by hand tally. Hand tallying with the aid of a stop watch, has been used in observing behaviour by recording who does what, with whom, when, where, under what conditions and for how long - all these variables being noted on card forms by means of codes, symbols, positioning and shorthand (Melbin 1954). Hand tallying obliges one to use arbitrary time units, even though these may be of short duration; units as short as three seconds have been used along with hand tallying (Moskowitz 1970).

Once the distribution of the frequency and duration of the event has been recorded and computed it is often necessary to re-group and interpret the results in terms of certain indices.

1.3 *The Problem of Interpretation*

Interpretive indices may be independent of the time record. They may depend on practical and theoretical considerations in the analysis of the subject matter analysed.

According to his interaction theory of personality, for example, Chapple interpreted total durations of interruptions minus the silences as an index of adjustment (Matarazzo 1956). Flanders, on the other hand, interpreted his data in terms of interaction matrices (Flanders 1970).

Such indices, matrices and other numerical interpretations permit one to classify events and their participants into functional categories.

1.4 The Problem of Categorization

Categories often depend on correlation with other phenomena outside the event in question. For example, it may have been discovered through experience or experimentation that people with dominant personalities make poor servants but good masters, or that teachers whose lesson matrices show a lot of student initiative are successful at teaching certain subjects to certain age-groups. The activity of such teachers could be put into the category of 'indirect teaching' (Flanders 1970). Similarly, by using personality indices, persons may be classed as good sales types, supervisor types - or even as neurotics.

The technology of visual and sound recording has become advanced and widespread enough to make the recording of all sorts of events practicable. Later analysis, interpretation and categorization of the components of such events depend on the purpose, function and state of knowledge and experimentation of the field observed. Computation of time variables, however, depends on advances in instrumentation. This is the aspect of the problem to which we shall now give our attention.

2. The solution

2.1 *Instrumentation*

One might say that the history of instrumentation in this field begins with the invention of the stop-watch, which was and still is widely used in time-and-motion studies and in games of competition. In fact this it what many people interested in chronometric variables, including the writers, sarted using, until it became evident that such a simple instrument could not meet the complex requirements in the analysis of multiple time variables in behaviour.

In the mid-1930s, Eliot D. Chapple, who was then developing his interaction theory of personality, was faced with the problem of measuring the dynamics of personal relations; and he attempted to do so through a study of the duration and frequency of personal contacts, changes in which, he noted, were always associated with expressions of changes in emotion and attitude. Using a simple stop watch, he recorded the duration of contacts in interviews between pairs of individuals and plotted the results on paper having ruled divisions, each equivalent to one minute.

Some time later, this procedure, which was burdensome, limited and imprecise, was replaced by the use of a recording device invented for the purpose. This machine included a paper tape, made to advance at a steady rate of speed, on which a stylus drew a line each time a key was pressed. With one such device per person it was possible to obtain a graphic record of his total talking (by adding the lengths of the lines), the number of times he started or stopped talking (by counting the beginnings or ends of the lines) and the length of his silences (by counting and measuring the intervening spaces between the lines). The analysis of all this was time consuming; about fifteen hours were required to analyse an interview of a half minute.

In the early 1940s, Chapple perfected a machine which would do most of the computing automatically. It was the first in a series of what became known as "interaction chronographs". On these, interviews were recorded by an observer using two keys; a computer, sometimes in a separate room, aided in producing a series of graphs at the end of the interview. After several years of use, however, more variables were needed in research than the first chronograph could measure. A perfected model was therefore developed by the inventor in the late 1940s. In this model, the observer's box was equipped with push buttons, two of which turned the recording machine on or off, two were allotted to the participants (one for the interviewer and the other for the interviewee) and another button activated a counter. After truning on this counter and the two power buttons, the observer kept one finger on the interviewer button and another finger on the interviewee button, pressing down or releasing when the subjects started or stopped talking. At the same time, a computer calculated the relative frequency and direction of the talking. The computer was designed to supply ten different measures by combining the frequency and duration of the stretches of speech and silences of the two persons. The first two measures (in the first two columns) gave the number of exchanges for each person by adding one unit each time one of the two interview buttons was pressed. Two columns indicated the number of times one of the persons had interrupted the other; two columns gave totals for the duration of the interruptions of each party, by subtracting the failures to respond; one column gave total speech or activity time of the interviewee; another added this to the silences, another gave the frequency of the silences (both buttons up), and finally a column indicated the frequency of the interruptions (both buttons down).

Each of these measures was given a certain interpretation. For example, duration of a person's speech (or action) became the activity index. The number one times he came out talking after an interruption became the dominance index (Chapple 1949).

Although used widely for the analysis of interviews, it must be pointed out that the interaction chronograph, in the final analysis, dealt with only two first-order variables - activity and non-activity. Only two people were dealt with at a time - the interviewer and the interviewee. The process was too complicated for some simpler purposes, the recording was not automatic, and the observer's reaction time and fatigue were sources of error.

In 1949, an attempt was made to eliminate the human observer of the interview - and at the same time reduce human error - by designing an automatic interaction analyser to record and compute the speech activity of an individual. The speech was first recorded by individual microphones - one for each person. Each microphone fed a separate tape recorder and computer. Duration and frequency wwer computed in intervals of 30 seconds; a counter and totalizer of duration gave the frequency distribution in these 30-second units, and finally there was a counter and classifier of interruptions. The apparatus consisted of electromagnetic counters, time delay relays, a stepping switch, and a pulse timer (Verzeano 1949). Although the machine was used in the field of psychiatry, for studies of free association, it was not universally applicable. The 30-second unit was arbitrary and too long for some studies, and a separate machine was needed for each person involved in an event.

In the 1950s a number of automatic devices came out for specific and limited operations in speech analysis. Some were designed only for the study of verbal activity during an interview (Kals & Mahl 1965). Others were invented for measuring rates of speech (Irwin & Becklund 1953) and variables in verbal conditioning (Krasner 1958).

In 1969, Starkweather conceived a speech-rate meter for verbal behavioural analysis. It produced a graphic cumulative record of pulses of speech, which showed a high correlation with the results of word-frequency counts obtained through typescripts (Starkweather 1950).

These counters of speech units are to be distinguished from other automatic speech analysers developed for phonetic research and language instruction. The latter goes back to Homer Dudley's vocoder of the year 1928. Much later, in the 1940s, the Bell Acoustic Spectograph was developed for the purpose of making speech 'visible' (Potter *et al* 1947). This later became the Sonograph, a version of which was used in the study of voiceprinting, each person having his characteristic print, presumably as inalterable as his fingerprints. These, along with the Mingograph, and the Eric, developed by Lucien LeBourhis at Laval University, were all used for the purpose of breaking down sound into its acoustic components.

Also in another category are more than a dozen types of speech recognition devices, like the SAID (Buiten & Lane 1965) developed under Harlan Lane in the mid-1960s to monitor practice in the learning of the intonation and rhythm of a foreign language. Here students record their speech on tape and look at a meter which displays the degree of acceptability of their utterance as far as the rhythm and intonation are concerned, after which the student re-shapes the intonation to come closer to the model imitated. Another, more elaborate device also using a small computer, analyses and recomposes an intonation curve on a television screen; it was developed by Pierre Léon in the late 1960s at the University of Toronto to permit students to compare their tone curve with the model.

Meanwhile researchers were still working to improve on the interaction chronograph. In 1965, a description was published of a new device called the interaction recorder which consisted of a timer and a punch unit. The data recorded was the same as that of the interaction chronograph; but it permitted faster processing of these data. It recorded on punched paper tape the sequences of interview speech and silence, keeping a record of when one or the other, both or neither of the persons was speaking. The tape was then fed into a Burroughs E101 computer which was programmed to produce a unit-by-unit analysis of the speech, silences and interruptions of the interview (Wiens 1965).

All these interaction recording devices have a great deal in common:

1. They deal with only three possibilities: activity, no activity or no types.
2. The activity or non-activity of only two persons is analysed.
3. From these two first-order measures, other second and third-order measures are derived.
4. They have narrow applicability, designed as they are mostly for interviews.

If other, more complex materials - lessons, games, television programs and motion picture films - were to be analysed, what was needed was a device that would record and compute variables in several types of activity and speech that can be used to analyse the simultaneous interaction of several people, a device that would give a greater number of first-order measures while being flexible and adaptable to different kinds of phenomena. This is what we attempted to do in developing the polychronometer.

2.2 *The Polychronometer*

It was with a view to meeting the needs for wider ranging instrumentation in the analysis of time variables in behaviour that the polychronometer was developed. In the context of what has been done in this field, it will now be easier to understand the history and development of this instrument, its design, mechanism, uses and special techniques.

2.2.1 History and Development

The development of the polychronometer began in the mid-1960s when one of the present writers attempted to relate teacher performance to method analysis (Mackey 1965a). The analysis of the teaching materials having been quantified (Mackey 1965b), it was necessary to find a way of quantifying their use in teaching performance. After attempting to use the chronometric techniques available at the time - all unsatisfactory - it was decided that the best long-term solution was to develop a machine for this and similar purposes. After a number of different designs and different types of instrumentation had been developed, a satisfactory working model was arrived at in late 1969, with the help of the chief technician of Laval University's Phonetics Research Laboratory, Lucien LeBourhis. After minor improvments and a few years of field testing, we were ready to report on the features, design, mechanism, uses and special techniques of this new device.

2.2.2 Main Features

The main features of the polychronometer are the following:

i) It is essentially a set of paired counter-time units each operated by a separate key.
ii) There are as many sets as there are fingers to operate them, that is, ten in number.
iii) This permits the analyst to make as many as twenty first-order measurements in a single run.
iv) Each key controls two types of measurement - duration and frequency.
v) Duration is measured in real time, since the electronic switches respond as quickly as an operator's reflexes.
vi) Each set is adaptable to automatic input (like speech recognition devices) or to computer output for first, second and third order measurement.
vii) The machine is light in weight and easily portable.
viii) Since the key-board is separate from the control box the machine can be operated at a distance or by remote control.

2.2.3 Design

The ten sets of timer-counters are imbedded in a sloped control box (18" x 12" x 6") containing the circuits and driving mechanism along with an on-off switch. The key board, with two sets of five concave keys, each designed for the varying length and shape of the fingers of each hand, is a separate unit which can be plugged into the control box - even at a distance. An optional remote control unit operates video-tape or magnetic tape playback machines for the analysis of recorded materials.

2.2.4 Mechanism

The mechanical problem in developing the machine was essentially one of circuitry and the coupling of appropriate relays to the activator keys (See Figure 1).

The electronic operation is described briefly as follows:

1. When one of the keys is touched, an activator establishes a + HT on a corresponding relay.
    1.1 The returning current passes through Q3 and R3, while Q3 remains active. (Base Q3 thru R3 and R4 to the HT, Q1 at rest.)
    1.2 Contact S1 auto starts the relay.
    1.3 Contact S2 discharge C2 on Base Q6 thru R1, which moves the corresponding key-touch counter ahead one unit, thru Q6.
    1.4 Contact S3 starts the corresponding timer.
    1.5 The tension developed by the Q3 emitter is not sufficient to start the monostable Q1-Q2, but sufficient to give to the base of Q5, thru R7, enough current to light the pilot light L1, indicating that one of the timers is active.
2. When another key is touched, its returning current, coupled with that of the previous one, produces enough tension thru R3 at Point C (1 volt) to start the monostable Q1-Q2, cutting the basic current Q3 (10 milliseconds). This unswitches the preceding relay while automatically putting a new relay into action behind the key which has just been touched. Anything over 10 milliseconds will thus be registered. The cycle is then repeated.
3. The stop-button switches off all active relays. It must be pressed immediately after the end of the period of analysis. Otherwise, the timer set in motion by the last key will keep on counting.

3. Polychronometric Techniques

The polychronometer can be used for the quantitative analysis of all phenomena that take place in time. It can measure the frequency and duration of ten categories - according to such criteria as place, person, object, organism, sound, action or speech - while the event is taking place, or after the event through a visual and/or sound recording of it.

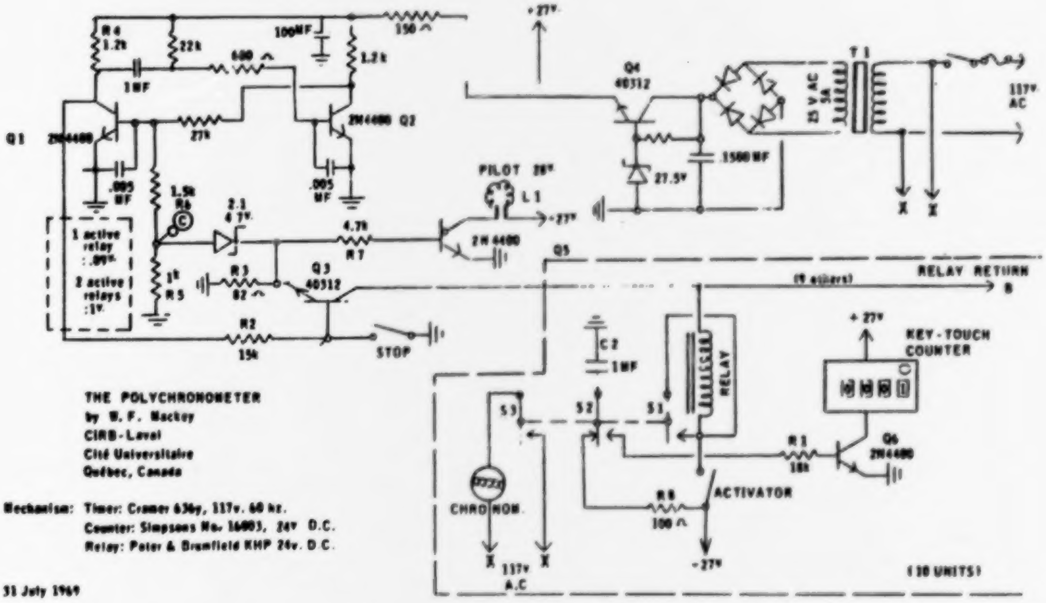
If more than ten categories have to be accounted for, it is necessary to use a recording of the event on film, videotape or magnetic sound tape. The categories can be divided into sets of ten, and as many re-runs made as required. For a thorough analysis of what can take place in a classroom lesson, for example, as many as three re-runs may be used.

Any special techniques needed for the efficient use of the polychronometer can be mastered in a few hours of practice. Efficiency of use can be increased if the tendencies of the observer-operator and some basic psychological truths are taken into account. Here are a few:

We divide the categories, if possible, according to a basic dichotomy (e.g. teacher-class, interviewer-interviewee, action-speech, left side-right side, etc.). One side of the dichotomy can be allotted to the right hand and the other to the left hand key board. We allot the most frequent or likely categories to the index and middle finger (e.g. teacher-talk). We avoid, if possible, the allocation of similar categories to adjacent keys. Techniques of usage have to be adapted to the type of material being observed and to the aims of the analysis.

FIGURE 1

The Polychronometer



In sum, there was a growing need for instrumentation which could enable us to observe and compute phenomena that take place in time. Although problems of observation, computation, interpretation and categorization varied from field to field and from problem to problem, it was possible to design an instrument for use in any situation where time-variables had to be measured as duration and frequency. The polychronometer was used for such purposes.

In the following chapters, polychronometric techniques are demonstrated in the context of lesson analysis for the study of the teaching-learning relationship, and the measurement of the effects of specific types of teaching behaviour on the learning of such language skills as speaking and reading. Finally there is a demonstration of the use of the techniques in measuring the efficiency of language teaching.

Chapter 2

POLYCHRONOMETRIC PROFILES OF TEACHING BEHAVIOUR

1. **Categorization**
    - 1.1 *Teaching Behaviour*
        - 1.1.1 Overall Behaviour
        - 1.1.2 Didactic Behaviour
    - 1.2 *Student Behaviour*
        - 1.2.1 Overall Behaviour
        - 1.2.2 Learning Behaviour
        - 1.2.3 Language Behaviour

2. **Computation**

3. **Interpretation**
    - 3.1 *Lesson Profiles*
    - 3.2 *Indices of Behaviour*

4. **Sample Lessons**
    - 4.1 *Polychronometric Profiles of Lesson A*
    - 4.2 *Polychronometric Profiles of Lesson B*
    - 4.3 *Polychronometric Profiles of Lesson C*

Key to Codes

Introduction

Our first objective in the use of polychronometric techniques is to demonstrate how they can be used to distinguish different types of teaching behaviour and teacher-oriented learning behaviour. We illustrate the techniques by establishing profiles for three sample language lessons, on the basis of which we give examples of comparative polychronometric analysis.

Like other techniques of behaviour analysis, polychronometric techniques incorporate three elements: categorization, computation and interpretation of behaviour.

1. Categorization

Some forty behavioural categories can account for the teacher's and students' didactic behaviour as well as the students' language performance (Table 1). Didactic behaviour, on the part of the teacher, refers to any technique used to teach the language or to manage the class; on the part of the students, it refers to the activities they engage in to learn or practice the language. The students' language performance is also analysed, because the extent to which they may be said to have learnt the language is determined not only by the activities they engage in, but by the language forms they produce.

Of the forty categories, fourteen designate teaching behaviour, twenty-four are for student behaviour and two for Silence. It will be observed that two categories, Silence and Group Activity, appear twice. This provides some measure of flexibility to the system. The users may be interested only in analysing the general characteristics of a lesson; in which case, they need use only the behaviours in the overall behaviour dimension. For a more detailed analysis, they would include the other two dimensions.

TABLE 1

Categories Polychronometric Technique of Foreign Language Teaching Analysis

Overall Behaviour

Teacher		Student	
1. Action 2. Action and Speech 3. Observation and Speech 4. Speech	10. Silence	5. Action 6. Action and Speech 7. Observation and Speech 8. Speech	Individual
		9. Group Activity	

Didactic Behaviour

11. Presentation
12. Confirmation
13. Rejection
14. Administration
14. Questioning
16. Prompting
17. Modeling
18. Reading
19. Writing
20. Use of the native language

Learning Behaviour

Individual	Correct	21. Reading aloud 22. Variational drills 23. Imitative and Incremental drills 24. Operational exercises
	Incorrect	25. Reading aloud 26. Variational drills 27. Imitative and Incremental drills 28. Operational exercises
29. Group Activity 30. Silence		

Language Behaviour

31. Stammering
32. Lexical error
33. Grammatical error
34. Phonetic error
35. Phonetic and lexical error
36. Phonetic and grammatical error
37. Grammatical and lexical error
38. Correct statement
39. Correct statement after correction
40. Incorrect statement

1.1 *Teaching Behaviour*

This includes overall or generalized behaviour and specific goal-oriented, in this case, didactic behaviour.

1.1.1 Overall Behaviour

Action: Any nonverbal behaviour directly related to the material being studied. This would include pointing to objects as a way of asking questions, presenting new material or reviewing known material; gestures and mimes.

Action and Speech: Speaking in the foreign or native language while performing appropriate actions, e.g. the teacher slowly closes a window while uttering the sentence *I am closing the window.*

Observation and Speech: Speaking in the foreign or native language while looking at an object or picture or any graphic representation (e.g. a flash card).

Speech: Teacher talk unaccompanied by specific action or observation.

1.1.2 Didactic Behaviour

Presentation: This includes presentation of new material or review of known material. The procedures used could be differential (explanations or translations), ostensive (use of objects, actions or situations), pictorial (use of pictures) or contextual (definition, enumeration, substitution, etc.). The time involved can range from that of a short utterance to a lengthy lecture. The teacher may be presenting language forms so that the students can repeat them, conveying information about material studied, explaining a grammar pattern or explaining the meaning of an item.

Confirmation: The teacher confirms or approvingly repeats student utterance. In some cases it may be slightly correctional; it does not however evoke an additional repetition. Included in this category are statements such as *Very good, That's Right, Of course it's snow.*

Rejection: The teacher rejects student response. Included here would be such statements as *That's wrong, No, this isn't snow*, or some form of nonverbal reaction such as a shaking of the head from side to side.

Administration: The teacher uses the target language in the mechanics of running the class.

Questioning: Included in this category are questions either personalised to individual students or relating to the material being studied. Typical examples would be questions such as *What's your name?, What do you call the baby cow?* The questions may be explicit or implicit, verbal or nonverbal. The teacher could ask a question intending that several students answer it. After the first response, the teacher may continue to ask the question of other students by merely raising the eyebrows with an enquiring look, ask a question of one student and continue questioning the same or other students by pointing to different objects. Implicit in the gesture is the question *What's that?* or *How would you describe that?* Also included here is the teacher's use of commands as a means of verifying students' comprehension of the target language, e.g. *Go to the door, Open the door.* Student response would usually be a physical one (cf. *The Total Physical Response Technique*). It is up to the observer to distinguish between such a command and one which is intended for classroom management.

Prompting: The teacher provides the students with language forms they have been momentarily unable to produce. The prompting is associated with encouragement and is not a correction of forms already uttered by the student.

Modeling: This category includes any utterance the teacher is modeling in the foreign language. It also includes the teacher's corrective representation of a student utterance.

Reading: Included in this category is any type of activity related to this skill; e.g. reading a narrative.

Writing: Writing on the blackboard, on transparencies for overhead projector, or any convenient surface.

Use of the native language: Included here are *all* uses of the native language, including use of target language words or phrases within native language frames.

1.2 *Student Behaviour*

This is also divided into general or overall behaviour, learning behaviour and language behaviour.

1.2.1 Overall Behaviour

Action: Any nonverbal behaviour directly related to the material being studied. This would include pointing at objects as a way of answering questions; gestures; mimes; carrying out directions to indicate listening comprehension.

Action and Speech: Speaking in the native or foreign language while performing appropriate actions. Dramatizations are included in this category.

Observation and Speech: Speaking in the foreign or native language while looking at an object, picture, or any graphic representation.

Speech: Student talk not accompanied by related action or observation.

Group Activity: Included in this category are all behaviours performed by two or more students simultaneously.

1.2.2 Learning Behaviour

Reading aloud: This category is self-explanatory.

Variational drills: This category includes student behaviour related to any one of the following exercises: conversion, transformation, substitution or re-statement of language forms; sentence completions; answers to multiple-choice type questions.

Imitative and Incremental drills: The students repeat language forms, model dialogues, songs, provides language forms in addition and expansion drills, and in sentence-building exercises.

Operational exercises: Included in this category are all types of creative behaviour on the part of students: answers to questions and look-and-say drills; descriptions, narrations, expositions and translations; physical responses.

Group activity: All behaviours performed by two or more students simultaneously.

Silence: Absence of any type of verbal behaviour or meaningful nonverbal behaviour. Meaningful nonverbal behaviour is any type of physical response related to the material being studied.

1.2.3 Language Behaviour

These behaviours are for the most part self-explanatory. Each student utterance is identified as belonging to one of those types. For example, the statement **Yes they does* is recorded as one grammatical error. An incorrect statement is any statement which the teacher rejects.

2. Computation

The forty behaviours to be identified are first divided into four sets of ten, one set for each of the behavioural dimensions to be analysed. The four sets represent:

A. Overall teacher-student behaviour
B. Teaching Behaviour
C. Learning Behaviour
D. Language Behaviour

Each set is then analysed in turn. In each set, one of the behaviours to be identified is assigned to each of the ten keys on the keyboard. Wherever possible, the allocation of behaviours to the right-hand or left-hand keyboard followed logical dichotomies: teacher/student in Set A, correct/incorrect in Set C. During recording, a key is activated the moment the behaviour to which it corresponds begins, and is reactivated only when that behaviour is completed. Cumulative frequency and relative duration of each behavioural category are thus recorded on the counter and timer units respectively.

3. Interpretation

By integrating the readings on the counters and timers for each behaviour component, profiles and indices of behaviour can be compiled.

3.1 *Lesson Profiles*

Two profiles are plotted for each lesson, one for frequency and another for duration. In addition, the relation between frequency and relative duration of the didactic behaviours of the teacher and the student is shown by plotting the frequency of each behaviour against its duration (See Figures 2 and 3).

The frequency and relative duration of each of the behaviours identified are expressed in percentages on a single decimal scale. The percentages can be arranged in a circle, around a circular percentage grid, so that the measurements of all the behaviours can be seen at a glance. All the plottings are connected by a continuous line to obtain a profile of the lesson as a whole (See Figure 4).

3.2 *Indices of Behaviour*

When the distribution of frequency and duration of the behaviours are recorded, computed and plotted on the lesson profile, the results are re-grouped and interpreted in terms of the following indices:

i. *The Activity Index:* This is a measure of the rate of interchange between teacher and students, or of the rapidity of the interaction. It is obtained by dividing total class time (expressed in seconds) by the sum of the frequencies of teacher and student didactic activities. Although ten different types of behaviour are isolated in the student sector, there are in fact only five different activities. A student who reads incorrectly is not engaging in a different activity from the one who does so correctly; therefore the correct/incorrect dichotomy is dropped when *type of activity* is being considered. The frequencies are thus combined: e.g.

Type of behaviour: Operational (correct) = 48
Operational (incorrect) = 12
Type of activity: Operational = 60

Any type of activity carried out by means of a piece of audiovisual equipment is also included. Silence, as it has been described above, is a state of inactivity. The higher the index, the less active the lesson.

ii. *The Variety Index:*

$$\frac{\text{Total class time in seconds}}{\text{Types of activities}}$$

The *Variety Index* is another indication of the type of interaction which takes place during the lesson. It differs from the *Activity Index* in that it is an indication of the average amount of time spent on each type of activity rather than the average duration of each individual activity. A very varied lesson is an active one, but an active lesson is not necessarily varied.

iii. *The Intensity Index:* This is a measure of the amount of pupil participation in the lesson, and it is an indication of the average number of student statements and actions per minute. It is calculated by dividing the number of individual and choral statements and actions by total class time expressed in minutes.

iv. *The Index of Effectiveness:* To indicate how effective a lesson is, student language performance is measured as a function of new teaching points introduced in the lesson. Correct sentences or phrases containing those points are token-counted and used in the following formula:

Index of Effectiveness:

$$\frac{\text{Total class time (seconds)}}{\text{Correct phrases}} \times \text{Number of new teaching points}$$

v. *The Degree of Structuration:* A lesson plan which could include one or more of the following phases is reconstructed from the original sequence of activities.

- a. Learner Preparation - Psychological
  - Linguistic
- b. Presentation
- c. Guidance - Comprehension
  - Production
- d. Habit Formation - Expression Drills
  - Content Drills
- e. Application - Conversation
  - Extensive reading
- f. Summing-up

The number and ordering of the phases become a measure of a lesson's degree of structuration.

4. Lesson Samples

The lesson samples analysed in this study were collected from three elementary school classes in the Quebec city region. In each case, English was the second language taught. The classes are identified as Class A, B and C.

Class A: Twenty eight Grade 3 students in their first year of English.
Class B: Thirty-one Grade 5 students.
Class C: Thirty Grade 5 students.

The lesson filmed in *Class A* was in the main a review lesson, while in *Classes B* and *C*, new material was presented. *Class B* studied the material in Block 4 Unit 23 of *Introducing English - One*, while *Class C* studied Units 23 and 24. No audio-visual equipment was used in *Class A*, but in *Classes B* and *C*, a tape-recorder was used to present the recorded texts of Units 23 and 24.

The three lessons were videotaped and audiotaped. The videotape recordings were made in collaboration with the Audio-Visual Department of Laval University. Their technicians made a technical survey prior to each videotaping session to decide on the positioning of cameras and microphones. On the day of the filming, the microphones and two cameras were set up in the classroom before the lesson began. The videotape recorder and the monitor were set up in an adjoining room for lesson A, and in an outside corridors for Lessons B and C. During the Lesson, only the cameramen remained in the classroom. The operators controlled and supervised the filming from outside.

It is this set of three recorded language lessons that was analysed according to the above polychronometric techniques to yield the different profiles (A, B and C) with corresponding indices of behaviour which are illustrated below.

4.1 *Polychronometric Profile of Lesson A*

4.1.1 Overall Behaviour

The cumulative counters on the polychronometer totaled duration of the lesson at about 1079 seconds, or approximately 18 minutes. Individual nonverbal communication took up 8% of total class time; 1.8% representing teacher action, and 6.2%, student action. Total teacher behaviour took up 57.9% of class time, student behaviour, 39.5%, with periods of silence accounting for a further 2.4%. The percentage figures for the cumulative frequency and duration of teacher behaviours on the one hand (Fr. 55.7, Dur. 57.9), and student behaviour on the other (Fr. 39.9, Dur. 39.5) would seem to indicate that there was relative uniformity in the length of each statement or action. While this may apply to teacher behaviour, it is not true of that of the students. The figures in Table 2 show that while Group Activity was the second largest time consumer, it was among the least frequent of class behaviours.

TABLE 2

Teacher-Student Behaviour in Lesson A

Teacher								Student							
Ta		Tas		Ts		Sl		Gr		Ss		Sos		Sa	
1.8%	4.5%	7%	6.7%	49.1%	44.7%	2.4%	4.3%	25.3%	8.6%	5%	13.7%	3%	8.2%	6.2%	9.4%
57.9%								39.5%							
55.7%								39.9%							

☐ Duration

▨ Frequency

Ta - Teacher Action, Tas - Teacher Action and Speech,
Ts - Teacher Speech, Sl - Silence, Gr - Group Activity,
Ss - Student Speech, Sos - Student Observation and Speech,
Sa - Student Action.

4.1.2 Teaching Behaviour

That Lesson A was mainly a review lesson is reflected in the low frequencies of behaviours concerned with such activities as presentation (4.3%), modeling or correcting language forms (2.9%), prompting (.9%), and rejection of student responses (.9%). The percentage figures indicating the duration of these activities are correspondingly low: 10.2%, 2.2%, .9%, and 1.1% respectively. Questioning and confirmation of student contributions, on the other hand, were very frequent, and together they took up 64.7% of teacher time. Figures 2a and 2b show the distribution of each of the behaviours identified in terms of their frequency of occurrence and relative duration.

4.1.3 Learning Behaviour

The 156 seconds of uninterrupted choral singing at the beginning of the lesson is largely responsible for the frequency-duration configuration of "Group Activity" in Figure 3. This and several shorter spells of choral reading explain why group activity represents only 20.1% of all student behaviours and yet took up 61.3% of student time.

Correct operational behaviour by individual students presents quite a different picture. Here, they are frequent (62.3%) but brief (30.5%). The same tendency towards brevity is apparent in their incorrect operational behaviour (Fr. 9.1%, Dur. 3.8%), and their imitations of teacher statements (Fr. 1.8%, Dur. .4%). In fact, each individual student behaviour, including their silent moments, lasted on an average only 1.8 seconds.

4.1.4 Language Behaviour

It was difficult to analyse the language behaviour of the students when they performed as a group. The choral reading, in particular, was slow, laboured and disjointed. Some students would begin reading a sentence before others did, so that individual word boundaries seemed to disappear. For this reason, only individual utterances were analysed.

Of the 56 utterances recorded, 46 or 82.1% were correctly produced at the first attempt. Phonetic mistakes were the most frequent; they accounted for all but one incorrect statement (See Table 3).

4.1.5 Profiles of Lesson A

For purposes of interpretation, the results obtained from the analysis of Lesson A were presented separately. The Behaviours identified are now brought together and displayed first in terms of their relative duration (Figure 4) and then according to their frequency of occurrence (Figure 5).

4.1.6 Indices of Lesson A

Lesson A was an active as well as a varied one. The particular nature of student group activity however, made it difficult to establish the true indices of Intensity and Effectiveness. There were no teaching points as such, but rather teaching areas, e.g., naming colours, fruits, vegetables. There were in addition, 27 instances of nonverbal student activity.

There was not much structuration in this lesson. For each teaching area, the teacher oriented the students by telling them what to expect and what kind of an answer to give (Ex. *I will give you a colour in French and you will give me the name in English*); she then asked the questions, and confirmed or corrected the students' replies. The sequence of activities was as follows: (For key to code, see end of chapter on pages 45 and 46).

1. O - (Q - A - Con)^n
2. O - (Q - A - Con)^n
3. O - (Q - A - Con/Corr)^n
4. O - (Q - A - Con)^n - P - Q - A - Con
5. O - (Q - A - Con) - Q - P
6. O - (Q - A - Con)^n
7. O - (Q - A - Con)^n
8. O - (Q - A - Con)^n - P - R
9. O - (D - A$_{ph}$)^n
10. (Q - A - Con)^n

O (Orientation, Q (Teacher Questioning), Con (Confirmation), Corr (Correction), P (Presentation), R (Repetition), D (Direction), A (Student Answer), *ph* (Student Answer, Physical), *n* (Frequency, 1+, /- or.

4.2 *Polychronometric Profile of Lesson B*

4.2.1 Overall Behaviour

Lesson B was found to have lasted 1317 seconds, or approximately 22 minutes. During the lesson, a tape recorder was used for 144 seconds, or 10.9% of total class time, to present language forms that the students could imitate. Total teacher talk accounted for 70.7% of the remaining time, with student talk taking up 25.0%, and periods of silence, 4.3%. There was no purely nonverbal behaviour either by the teacher or the students.

Although so much of total class time was taken up by the teacher, the students actually spoke more often than he did. Of the total number of behaviours recorded, 49.8% were performed by the students, 47.4% by the teacher, with 2.8% representing pauses in the interaction. The percentage figures for the frequency and relative duration of each of the categories in this dimension are given in Table 5.

The class spoke more often than the teacher did because their responses were not all elicited by teacher statements; there was quite a lot of choral repetition of forms modeled on the tape recorder. While this did in fact affect the scores, a comparison of the figures still indicates that, on the average, teacher statements tended to be long, and student statements brief.

FIGURE 2a AND 2b

Frequency and Relative Duration of Teaching Behaviour in Lesson A

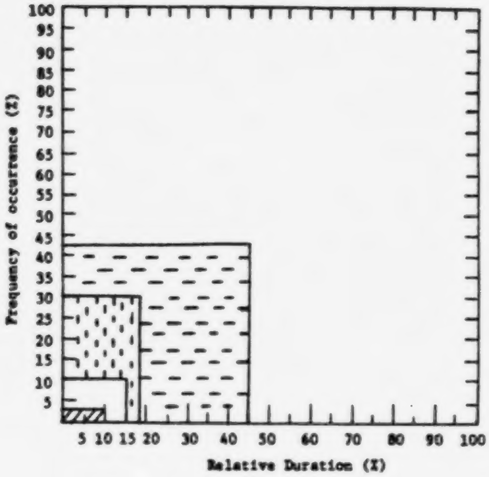


Fig. 2a

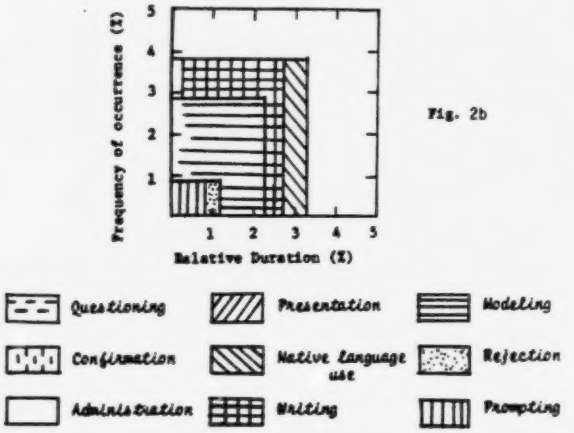


Fig. 2b

FIGURE 3

Frequency and Relative Duration of Learning Behaviour in Lesson A

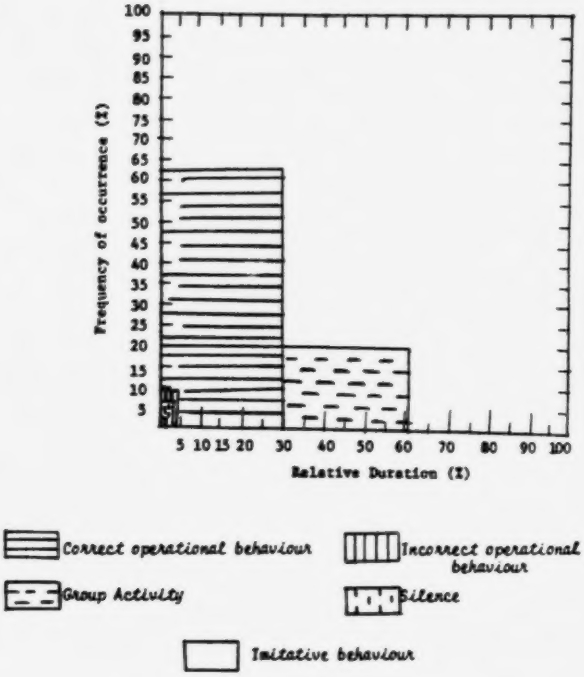


TABLE 3

Language Behaviour in Lesson A

	Le		Pe		Cs		Cc	
	1.4%	1.8%	19.7%	14.3%	76%	82.1%	2.8%	1.8%
Total					78.8%			
					83.9%			

☐ Duration

▨ Frequency

Le - Lexical Error, Pe - Phonetic Error
Cs - Correct Statement, Cc - Correct
Statement after Correction.

FIGURE 4

Duration Profile of Lesson A

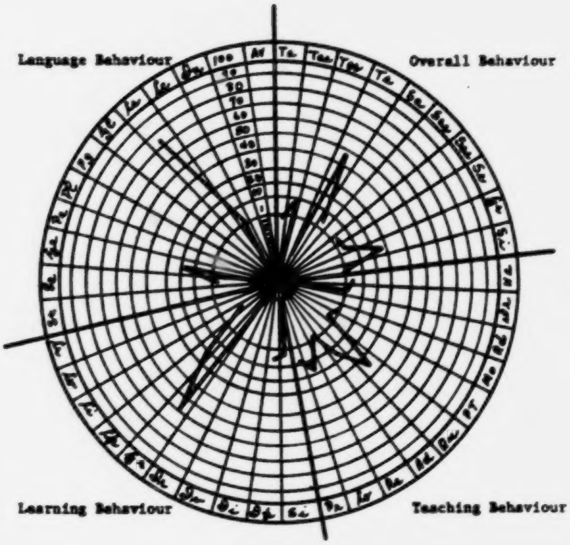


FIGURE 5

Frequency Profile of Lesson A

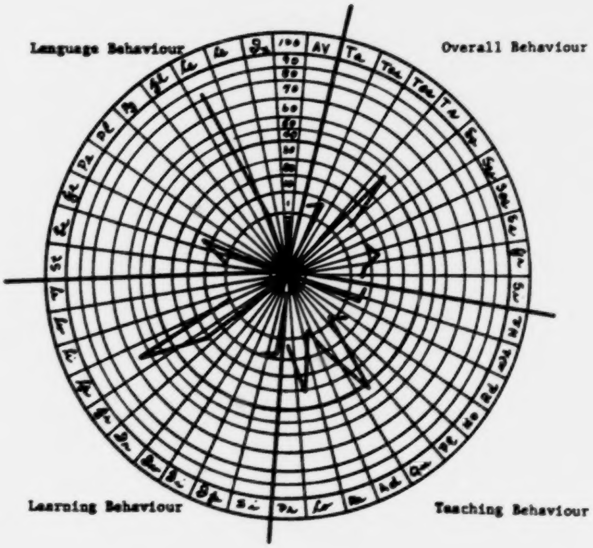


TABLE 4

Activities: Types and Occurrences: Lesson A

	Types	Occurrences
Teacher	Presentation	9
	Confirmation	64
	Modeling	6
	Prompting	2
	Questioning	88
	Writing	8
	Administration	21
	Use of the native language	8
	Rejection	2
Student	Imitative exercises	2
	Operational exercises	78
	Group Activity	22
Total		310

Activity Index $\frac{1079}{310} = 3.5$

Variety Index $\frac{1079}{12} = 90$

Intensity Index $\frac{60 \times 107}{1079} = 5.9$

Index of Effectiveness $\frac{1079}{60} \times \frac{1}{44} \times 7 = 2.9$

N.B. Numbers have been rounded.

TABLE 5

Teacher-Student Behaviour in Lesson B

	Teacher								Student					
	Tas		*Tos*		*Ts*		*Si*		*Gr*		*Ss*		*Sos*	
	10.6%	5.9%	28.2%	13.2%	31.9%	28.5%	4.3%	2.8%	10.6%	17.8%	2.5%	8.2%	11.9%	23.8
Total	70.7%								25.0%					
	47.6%								49.8%					

☐ Duration

▨ Frequency

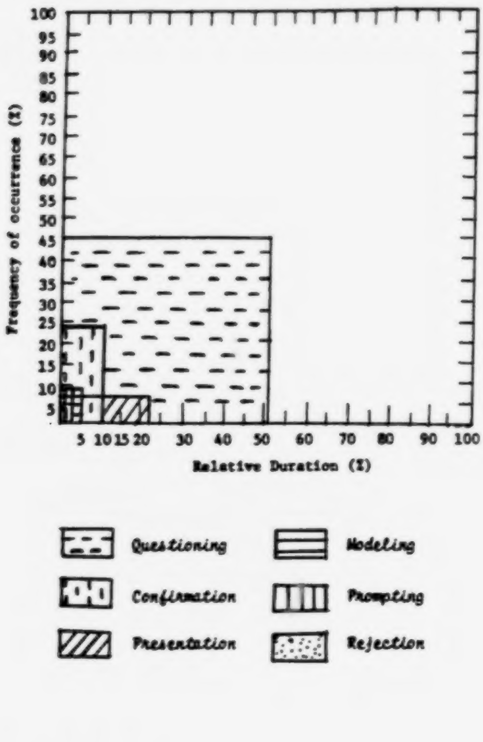
Tas - Teacher Action and Speech, Tos - Teacher Observation and Speech, Ts - Teacher Speech, Si - Silence, Gr - Group Activity, Ss - Student Speech, Sos - Student Observation and Speech.

4.2.2 Teaching Behaviour

The teacher spent more than half (56.9%) of his time asking questions of the students, but the actual number of questions he asked represented less than half of his total behaviour (45.9%). This would seem to indicate that he spent some time formulating them. In fact, he did tend to repeat a question twice, three times and sometimes even four times before calling on a student for an answer. He devoted 22.6% of his time to the presentation of the material being studied, but confined it to just a few long spells (Frequency 6.7%). Confirmation of student utterances were frequent but brief (Frequency, 23.9%), Duration, (10.4%). Only once did he use the target language in the mechanics of running his class, and then for only 7 seconds. On several occasions he either encouraged the students to continue or else provided them with language forms they were momentarily unable to produce. These however, seemed to have been for the most part, short, possibly one-word utterances (Prompting: Fr., 9.1%, Dur. 2.5%). He was more inclined to deal with incorrect answers by giving them the correct answer, than by rejecting their contributions outright (Modeling, Frequency, 9.7%; Rejection Frequency, 4.1%). Figure 6 shows the relation between the frequency and relative duration of each of the teacher categories. The dominant activity both in terms of frequency and duration was undoubtedly teacher questioning. The next most frequent activity was the confirmation of student contributions, but this took up much less time than his presentation of the material studied.

FIGURE 6

Frequency and Relative Duration of Teaching Behaviour in Lesson B

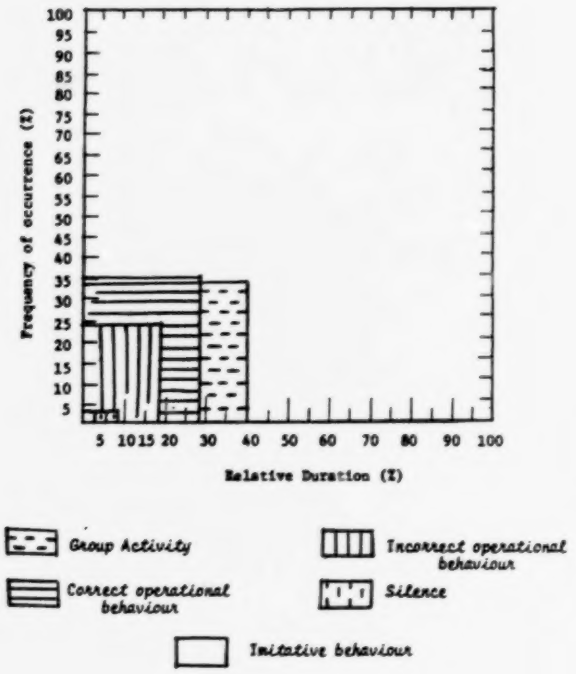


4.2.3 Learning Behaviour

We have already noted that student statements tended to be brief. In fact, each lasted on an average only 1.7 seconds (309 -÷184). Choral responses took up 40.1% of student time and represented 34.8% of total student activity. Individual student behaviour was predominantly operational. It made up 60.2% of all activity and took up 47.9% of student time. Only 35.8% of these however, were correct, accounting for 28.5% of the time allotted the students. There was not much repetition by individual students. Whatever little there was, (Duration, 2.3% Frequency, 2.2%) was correct. There were a few periods of silence. Figure 7 provides a graphic picture of the relation between the frequency and relative duration of student behaviour. Viewed in its entirety, no one type of behaviour seemed to have dominated. The almost square plottings on the grid would seem to indicate that the length of student statements remained almost constant throughout the lesson.

FIGURE 7

Frequency and Relative Duration of Learning Behaviour in Lesson B



4.2.4 Language Behaviour

The students, both chorally and individually spoke approximately 179 times during the lesson; 74.8% of the statements uttered were correct and these took up an estimated 78.3% of student talking time. Grammatical errors were the most frequent, accounting for 10.1% of all utterances. Table 6 shows the figures for the frequency and duration for each type of behaviour identified.

4.2.5 Profiles of Lesson B

An overall view of all the behaviours identified is shown in Figures 8 and 9. Figure 8 represents them in terms of their relative duration while Figure 9 displays them in terms of their frequency of occurrence. The differences in the lengths of the plottings indicate the type of relation which exists between the two variables in each behavioural dimension.

4.2.6 Indices of Lesson B

The 2.9 *Activity Index* established for this lesson indicates that it was an active one, with a change occurring on an average every three seconds. The ten different types of activities that the teacher and students engaged in gave the lesson a variety index of 131.7 (Table 7).

The amount of student participation has already been established; 49.8% of the behaviours recorded were performed by the students. The distribution of this participation is indicated by the *Intensity Index* which stands at 8.1; thus for every minute of class time, there were approximately 8.1 student utterances. Four new teaching points were introduced in the lesson. Seventy-eight sentences each containing one or other of those teaching points were correctly produced by the students; 39 were choral responses and 39 were individual responses. The lesson's *Index of Effectiveness* stands therefore at 1.12.

Index of Effectiveness: $\frac{1317}{60} \times \frac{1}{78} \times 4 = 1.12$

Degree of structuration

The original sequence of activities for the entire lesson can be summarised as follows:

Teaching Point 1:	$(P_{av} - R)^n - P_t)\ 2$ $(- A - Con/Corr)^n$
Teaching Point 2:	$(P_{av} - R)^n - P_t$ $(Q - A - Con/Corr)^n$
Teaching Point 3:	$(P_{av} - R)^n$ $(Q - A - Con/Corr)^n$
Teaching Point 4:	$(P_{av} - R)^n - P_t$ $(Q - A - Con/Corr)^n$ $(P_{av} - R)^n$ $(Q - A - Con/Corr)^n$
Summing-up:	$(Q - A - Con/Corr)^n$

P (Presentation), *av* (Tape Recorder), *t* (Teacher), R (Group Repetition), Q (Questioning), A (Individual Student Answer), Con (Teacher Confirmation), Corr (Teacher Correction), *n* (Frequency 1+, /-or).

There were therefore three phases to this lesson: the presentation and guidance of each of the teaching points, followed by a section devoted to the recapitulation of the items taught. The plan which was drawn up for this lesson might have read:

A1 - Presentation: Teaching Point 1
B1 - Guidance/Production: Teaching Point 1
A2 - Presentation: Teaching Point 2
B2 - Guidance/Production: Teaching Point 2
A3 - Presentation: Teaching Point 3
B3 - Guidance/Production: Teaching Point 3
A4 - Presentation: Teaching Point 4
B4 - Guidance/Production: Teaching Point 4
C - Summing-up

TABLE 6

Language Behaviour in Lesson B

	St		Le		Ge		Pe		Pg		Cs		Cc		In	
	4.3%	3.8%	1.4%	2.6%	7.6%	10.1%	.7%	.5%	1.1%	1.1%	76.1%	71.5%	2.2%	3.3%	6.3%	7.8%
Total											78.3%					
											74.8%					

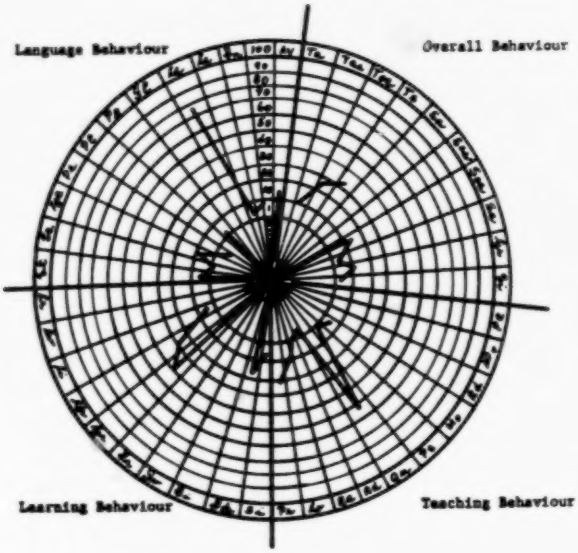
☐ Duration

▨ Frequency

St - Stammering, Le - Lexical error, Ge - Grammatical error, Pe - Phonetic error, Pg - Phonetic and grammatical error, Cs - Correct statement, Cc - Correct Statement after correction, In - Incorrect statement.

FIGURE 8

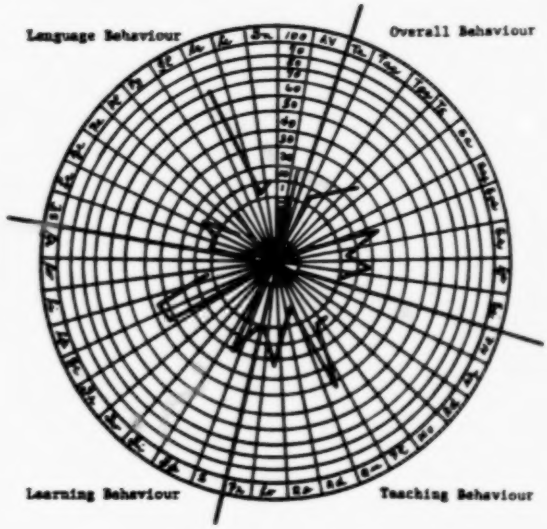
Duration Profile of Lesson B



N.B. *See key to code used in Profile*

FIGURE 9

Frequency Profile of Lesson B



N.B. *See key to code used in Profile*

TABLE 7

Activities: Types and Occurrences: Lesson B

	Type	Occurrences
Teacher/AV	Presentation	78
	Confirmation	47
	Modeling	19
	Prompting	18
	Questioning	90
	Administration	1
	Rejection	8
Student	Imitative exercices	4
	Operational exercices	111
	Group. Activity	64
Total		440

Activity Index: $\frac{1317}{440} = 2.9$

Variety Index: $\frac{1317}{10} = 131.7$

4.3 *Polychronometric Profile of Lesson C*

4.3.1 Overall Behaviour

The relative duration of this lesson came to about 1068 seconds or approximately 17 3/4 minutes. During that time, a tape-recorder was used for about 265 seconds, leaving 803 seconds for teacher-pupil interaction. Further analysis showed that 51.7% of that time was taken up by the teacher, 43.7% by the student, and 4.4% by periods of silence. But although teacher talk took up more time than student talk, the class spoke more than twice as often as the teacher did.

It is true that not all students' utterances were in reply to teacher statements, but the figures do indicate that while one teacher statement lasted on an average 5.3 seconds, only an average 1.7 seconds was spent on each student utterance.

4.3.2 Teaching Behaviour

A preliminary analysis of the teacher's behaviour indicated that her statements on the average tended to be long when compared to those of the students. However, a more detailed analysis showed that this tendency was confined to questioning and presentation which took up 75.3% of teacher time and yet represented 52% of the total number of teacher behaviours recorded. Questioning alone took up 58.4% of her time and accounted for 42.8% on her behaviour. Confirmation of student contributions was relatively frequent (16.3%), but brief (6.4%). They seemed to have been for the most part repetitions of student statements, for the average duration of each confirmatory remark is the same as that of a student statement: 1.7 seconds. The teacher modeled language forms fairly frequently (19.4%), and on a few occasions, provided students with forms they were unable to produce (9.2%). Figure 10 shows the relation between the frequency and relative duration of each of her behaviours. It is evident from the plottings on the grid that questioning was by far the most dominant, both in terms of frequency and duration.

4.3.3 Learning Behaviour

Choral responses dominated all others, accounting for 72.9% of all student behaviours recorded, and taking up 71.9% of student time (Figure 11).

Individual behaviour was largely operational (Duration, 17.3%, Frequency, 16.4%). The others were variational (Dur., 4.6%, Fr., 6.6%) and imitative (Dur., 1.9%, Fr., 1.6%) with periods of silence lasting 4.3% of the time (Figure 12).

TABLE 8

Teacher-Student Behaviour in Lesson C

	Teacher								Student							
	Tas		Tos		Ts		Si		Gr		Ss		Sos		Sa	
	5.9%	1.8%	32.7%	12.1%	23.1%	14.7%	4.4%	2.9%	33.2%	45.5%	2.1%	5.2%	7.7%	13.2%	.7%	.4%
Total	51.7%								43.7%							
	28.5%								68.1%							

☐ Duration

▒ Frequency

Tas - Teacher Action and Speech, Tos - Teacher Observation and Speech, Ts - Teacher Speech, Si - Silence, Gr - Group Activity, Ss - Student Speech, Sos - Student Observation and Speech, Sa - Student Action.

FIGURE 10

Frequency and Relative Duration of Teaching Behaviour in Lesson C

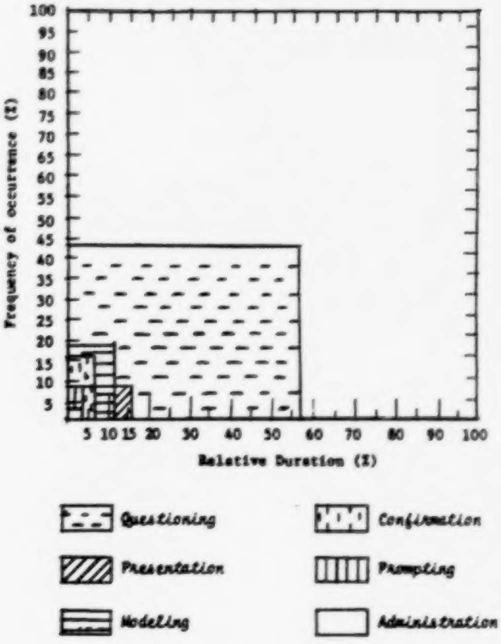


FIGURE 11

Group Activity in Lesson C

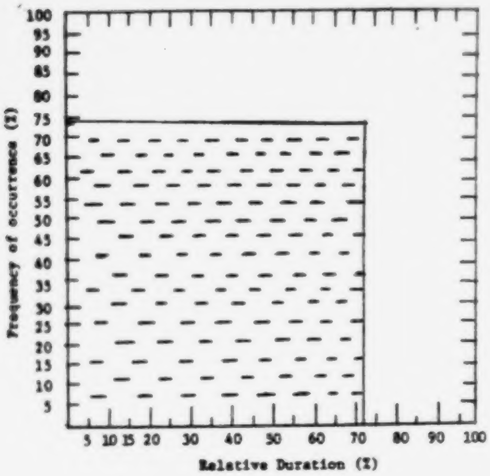
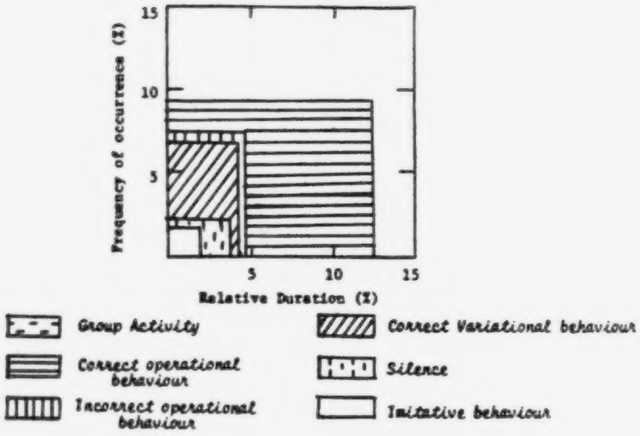


FIGURE 12

Individual Student Learning Behaviour in Lesson C



4.3.4 Language Behaviour

An analysis of the choral and individual behaviour of the students reveals that as many as 91.8% of all utterances were correctly produced; 1.6% after correction by the teacher (Table 9). There were as many grammatical errors as there were lexical ones, although the phrases in which the latter occurred tended to be shorter.

TABLE 9

Language Behaviour in Lesson C

	St		Le		Ge		Cs		Cc	
	1.1%	.5%	2.5%	3.8%	4%	3.8%	90.5%	90.2%	1.7%	1.6%
Total							92.2%			
							91.8%			

□ Duration

▨ Frequency

St - Stammering, Le - Lexical Error, Ge - Grammatical error, Cs - Correct Statement, Cc - Correct Statement after correction.

4.3.5 Profiles of Lesson C

This four-dimensional analysis has presented a rather fragmentary picture of what took place during the lesson. It is possible however, to get at a glance, a complete profile of the lesson as a whole in terms of the relative duration of the behaviours (Figure 13), and the frequency of their occurrence (Figure 14).

4.3.6 Indices of Lesson C

The interchange between the teacher and the students was a rapid one, as the *Activity Index* of 2.7 seems to suggest.

Activity Index: $\frac{1068}{398} = 2.7$

Variety Index: $\frac{1068}{10} = 106.8$

Student behaviour took up 43.7% of class time, with sentences or phrases being produced at the rate of 9.9 per minute.

Intensity Index: $\frac{60 \times 177}{1068} = 9.9$

Of the 164 correct utterances produced, 68 contained one or other of the four teaching points introduced in the lesson. As many as 60 of these were choral responses, with only eight being produced by individual students. The *Index of Effectiveness* stands therefore at 1.05.

Index of Effectiveness: $\frac{1068}{60} \times \frac{1}{68} \times 4 = 1.05$

The new items taught were all structural. They were presented in pairs, tried out, and finally revised. The original sequence is as follows:

Teaching Point 1: $(P_{av} - R)^{n} - (P_{t} - R)^{n}$
Teaching Point 2: $(P_{av} - R)^{n}$
$(Q - A - Con/Corr)^{n}$
Teaching Point 3: $(P_{av} - R)^{n} - P_{t}$
Teaching Point 4: $(P_{av} - R)^{n} - P_{t}$
$(Q - A - Con/Corr)^{n}$
Summing-up: $(P_{av} - R)^{n} - (Q - A - Con/Corr)^{n}$

P (Presentation), *av* (Tape-recorder, *t* (Teacher), R (Choral Response), Q (Questioning by Teacher), A (Individual Student Answer), Con (Teacher Confirmation), Corr (Teacher Correction), *n* (Frequency 1+, / - or).

A plan for this lesson therefore could have read:

A1 - Presentation: Teaching Point 1
A2 - Presentation: Teaching Point 2
B1 - Guidance/Production: Teaching Points 1 and 2
A3 - Presentation: Teaching Point 3
A4 - Presentation: Teaching Point 4
B2 - Guidance/Production: Teaching Points 3 and 4
C - Summing-up

The results obtained for each lesson and each system have so far been presented separately with no comparisons made between lessons or systems. This fulfills the first stage of the analysis: the isolation, identification and quantification of teaching behaviour in a second language classroom.

We are now in a position to compare these lessons along several significant dimensions.

TABLE 10

Activities - Types and Occurrences: Lesson C

	Types	Occurrences
Teacher/AV	Presentation	132
	Confirmation	16
	Modeling	19
	Questioning	42
	Administration	3
	Prompting	9
Student	Variational exercises	12
	Imitative exercises	3
	Operational exercises	30
	Group Activity	132
Total		398

FIGURE 13

Duration Profile of Lesson C

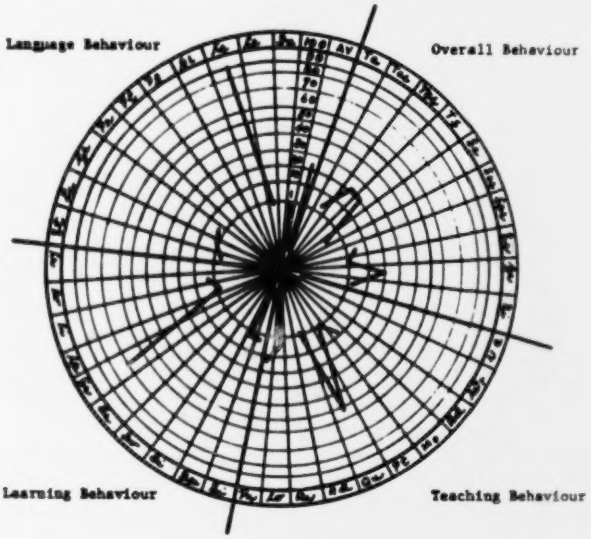
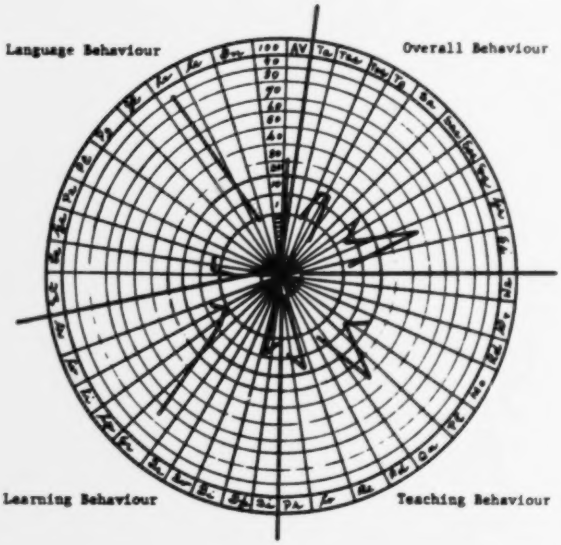


FIGURE 14

Frequency Profile of Lesson C



4.4 *Comparative Polychronometric Analysis*

To begin with, are three lessons of the same type? The answer is no. Lesson A differs from Lessons B and C. The high percent figure representing teacher confirmation of student contributions, and the low figures obtained for presentation, modeling and prompting suggest that most of the material covered in Lesson A was not new to the students. The sequence of events in Lesson A is another indication of the differences which exist between it and Lessons B and C. How do these two compare? Although Lessons B and C were basically the same, there were considerable differences in the duration and frequency of both teacher and student behaviours (Figure 15 to 18).

By this technique, however, we were able to analyse student behaviour in great detail, collecting information not only on the amount of choral and individual behaviour, but also on the type of language acquisition activities and the learning outcomes of these activities.

The rapid teacher-student interchange is confirmed by the low *Activity Indices* (Table 11).

TABLE 11

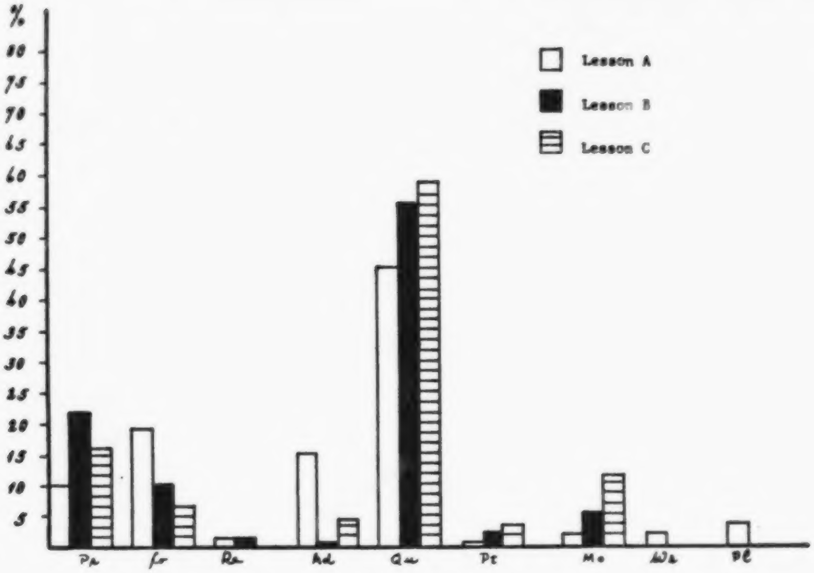
Polychronometric Indices for Lessons A, B and C

Indices	Lesson A	Lesson B	Lesson C
Activity Index	3.5	2.9	2.7
Variety Index	90	131.7	106.8
Intensity Index	5.9	8.1	9.9
Index of Effectiveness	2.9	1.12	1.05

N.B. Numbers have been rounded

FIGURE 15

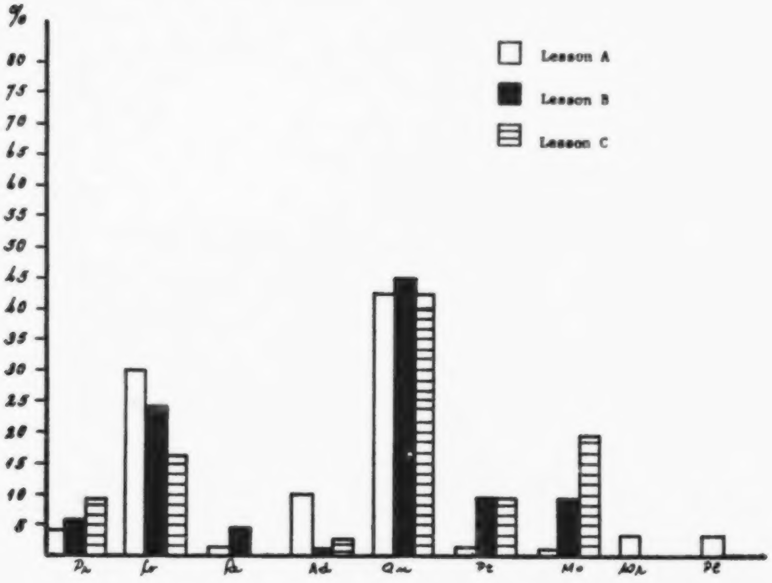
Distribution of Teaching Behaviours in Lessons A, B and C: Duration



N.B. *See key to letter codes used in histogram*

FIGURE 16

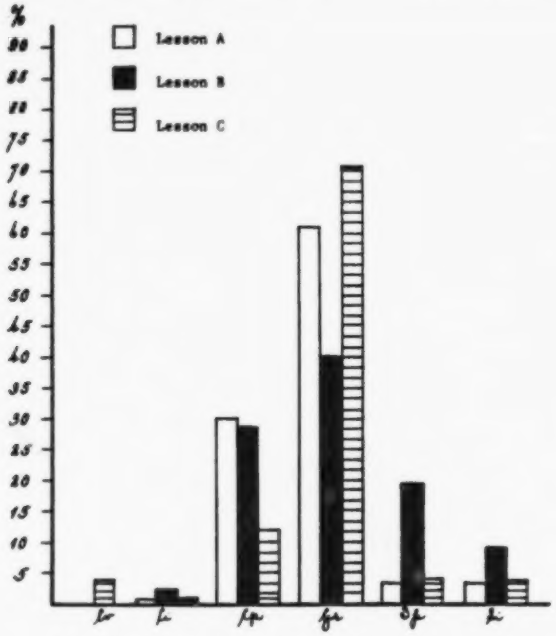
Distribution of Teaching Behaviours in Lessons A, B and C: Frequency



N. B. *See key to letter codes used in histogram*

FIGURE 17

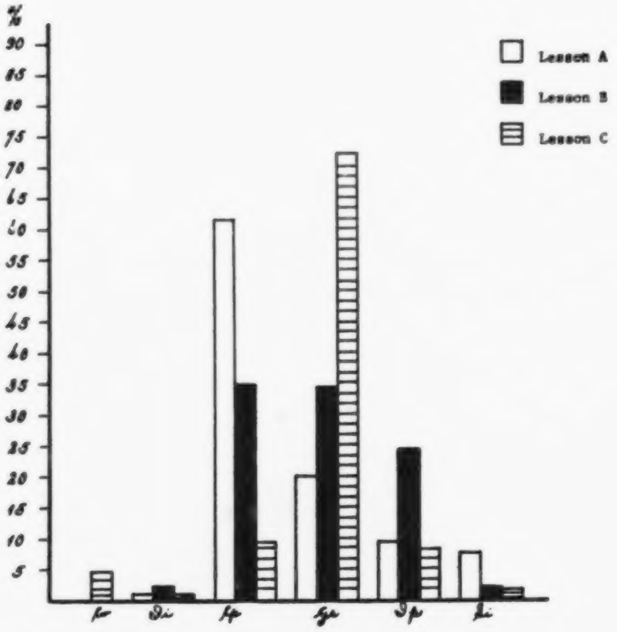
Distribution of Learning Behaviours in Lessons A, B and C: Duration



N.B. *See key to letter codes used in histogram*

FIGURE 18

Distribution of Learning Behaviours in Lessons A, B and C: Frequency



N.B. *See key to letter codes used in histogram*

The indices of intensity and effectiveness in Lesson A include a type of choral activity which does not lend itself to individual analysis. Lesson C, more active than Lesson B, was also more intense, but less effective. Of the 134 correct utterances produced by the students in Lesson B, 78 contained one or other of the four new teaching points introduced in the lesson: 39 were choral responses, and 39 were individual contributions. In Lesson C, 68 of the 164 correct student statements contained one or other of the four teaching points, but 60 of them were choral responses and only 8 represented individual contributions. There seemed to be at least two reasons for this:

1. The teacher in Class C did not insist on students' use of the teaching points in the formulation of their answers. Excerpts from the two lessons show the difference in interrogation techniques.

Teaching Point - *Has*

Lesson B
Teacher: What does he have, François?
François: It is
Teacher: No, not "it is"
François: He has
Teacher: He has
François: The red basket
Teacher: A red basket

Lesson C
Teacher: What does he have?
Pupil: A - a red shirt
Teacher: A red shirt and
Pupil: Basket, red basket
Teacher: O.K.

Teaching Point - *Does*

Lesson B
Teacher: Does he have a blue shirt?
Pupil: Yes, he does
Teacher: Yes, he does. Very good

Lesson C
Teacher: Does he have a red basket?
Pupil: No, he is not
Teacher: Listen. Does he have a red basket? Yes or No?
Pupil: No

2. Whereas each new teaching points in Lesson B was tested for comprehension and production before another was presented, the new teaching points in Lesson C were presented in pairs. This type of structuration did not allow enough room for a tryout of each new point separately. The students therefore, never quite mastered any of the new points introduced.

KEY TO CODE USED IN LESSON PROFILES

Overall Behaviour		*Teaching Behaviour*	
Ta -	Teacher Action	Pr -	Presentation
Tas -	Teacher Action and Speech	Co -	Confirmation
Tos -	Teacher Observation and Speech	Re -	Rejection

Ts	Teacher Speech
Sa	Student Action
Sas	Student Action and Speech
Sos	Student Observation and Speech
Ss	Student Speech
Gr	Group Activity
Si	Silence

Ad	Administration
Qu	Questioning
Pt	Prompting
Mo	Modeling
Rd	Reading
Wr	Writing
Nl	Native Language

Learning Behaviour

Cr	Reading aloud correctly
Cv	Correct variational exercise
Ci	Correct imitative exercise
Cp	Correct operational exercise
Gr	Group activity
Ir	Reading aloud incorrectly
Iv	Incorrect variational exercise
Ii	Incorrect imitative exercise
Ip	Incorrect operational exercise
Si	Silence

Language Behaviour

St	Stammering
Le	Lexical error
Ge	Grammatical error
Pe	Phonetic error
Pl	Phonetic and lexical error
Pg	Phonetic and grammatical error
Gl	Grammatical and lexical error
Cs	Correct statement
Cc	Correct statement after correction by the teacher
In	Incorrect statement

We have seen how it is possible to quantify differences in language lessons by measuring the length of each of their components along with their frequency of occurrence. We do not yet know, however, whether these differences are significant, and if so, which ones are the most important in producing language learning. This will be the concern of the next two chapters.

Chapter 3

LEARNING AND TEACHING BEHAVIOUR: COMPARING LESSON TYPES

1. **Procedures**
   1.1 *Collection of Data*
   1.2 *Data Analysis*

2. **Lesson Types**

3. **Behaviour Distribution Profiles**

To discover the extent to which differences in the components of a language lesson - their respective duration and frequency - are significant, it is necessary to measure and analyse the results of comparable lessons. It is possible to find language lessons that conform more than do others to the theoretical ideal, according to one of the several theories of language teaching. This, however, is no guarantee that the ideal gives better results. Is it possible therefore to find out the characteristics of lessons that do in fact produce results and the optimum proportion of the lesson components? In an attempt to answer these questions, we tested five classes each being taught the same lesson. The comparable groups were pre-tested on their knowledge of the content of the lessons to be taught. The actual lessons were videorecorded and analysed by polychronometric techniques similar to those described in the preceding chapter.

1. Procedures

To investigate the effectiveness of didactic behaviours (functional behaviours) we used the linguistic performance of 11-year-olds in beginning classes of English as a second language.

1.1 *Collection of the Data*

The learning variables used consisted of eleven concrete and two structure words. These words had been selected from an inventory of known and unknown words of the class text, *Introducing English*. The two grammar words were taken from chapter twenty-one and the concrete words were selected from chapters which were not yet taught, and from outside materials. Therefore, these variables were used in the construction of tests which measured linguistic performance; and they were taught in lessons by means of which didactic behaviours were studied. A trial test was administered to ten fifth graders. The results showed that learners had some language skills, or that they had used strategies which enabled them to score on the listening comprehension test. However, all students received zero scores on the speaking test.

Subjects from five elementary schools and five classrooms participated in the study project. The subjects - 119 in number - were all Grade V students and 11 years of age. And the data, which consisted of tests of listening comprehension and speaking, as well as lesson samples, had been collected from these five classrooms which were labelled: Lesson A, Lesson B, Lesson C, Lesson D and Lesson E.

1.2 *Data Analysis*

1.2.1 Statistical Analyses of the tests

The listening comprehension and the speaking tests were scored out of 100 points, and then submitted to Variance of Difference Scores and Covariance Analyses (the computer programmes SPSS and BMD04 were used for the calculations). The results obtained on both analyses were the same. On the listening comprehension test, there was a significant difference at the 5 percent level, and Scheffé's comparison of the mean scores revealed that Lesson A was different from Lesson E, but Lesson E = B = C = D. Although, it appeared that Lesson A was different from all the others, only the lesson analysis would show in what way this lesson was different from the four others. There was no significant difference, however, among the five classrooms on the speaking test. The mean gain scores had been used throughout the study, since they were simpler, and more readily understood than the more complex adjusted mean scores.

1.2.2 Polychronometric Lesson Analysis

The five lesson samples had been analysed separately by a polychronometric technique lesson analysis. The results of the quantitative analysis were plotted onto Behaviour Distribution Profiles. And the data from this first analysis were used in the calculation of behaviour indicators of activity, variety, intensity and performance.

1.2.3 Calculation of Behaviour Indicators

The data obtained on the quantitative analysis were used in order to calculate reliable quantitative indicators of activity, variety, intensity and performance through which the quality of the lesson could be assessed.

Therefore, the behaviour indicators have been calculated in the manner described below.

1. The activity indicators were found by dividing the cumulative frequency by the duration of the behaviour: Fr./Dr.
2. The variety indicators were found by dividing the sum of categories by the duration of the set: Category/Dr.
3. The intensity indicator was found by dividing the sum of actions and utterances by the duration of the lesson: actions and utter./Dr. (total class time)
4. Individual performance indicators were found in the following manner:
   - A (i) The sum of correct individual utterances divided by duration of set: Utter. (corr. ind)/Dr.
   - A (ii) This indicator was found by dividing the correct utter. by total number of utter. (ind.): corr. utter. ind./Total utterances.
   - A (iii) This was calculated by dividing the sum of error type by the sum of utterances (ind.): Error Type/ of utterances (ind.)
   - A (iv) This indicator was found by dividing the sum of errors of vocabulary and grammar by the sum of utterances (ind.): errors of Vocabulary & grammar/ utterances (ind.).

2. Lesson Profiles and Indices

Applying these procedures to each of the five lessons (A, B, C, D, E) and corresponding classes the following characteristics were observed.

2.1 *Lesson A*

There were two distinctions made of action and speech activities in Lesson A. They were classified as Didactic Behaviour and Language Behaviour. Didactic Behaviour consisted in the teaching and learning processes, whereas, language behaviour concerned the manner in which language was used: Teacher's language use and learner's language performance. Didactic behaviours were divided into categories and sub-categories of teaching and learning. These categories were made up of sets of 10 behaviours. The analysis distinguished three major categories and three sub-categories of teaching but only one for learning.

The results of the analysis showed the new didactic points had been used 472 times, and the relative duration of these new materials, for the entire lesson was 25.3 minutes. The lesson as a whole showed that the duration and frequency of teaching and learning activities were almost equal. Articulation and verification had been the most frequently used and of longest duration, of the seven teaching behaviours. In addition, the data obtained on the first analysis were plotted onto Behaviour Distribution Profiles (See Figure 19 to 28). Furthermore, the behaviour indicators of the lesson were calculated from there data. The analysis showed that Lesson A was as active as it was varied. However, the highest activity was observed during speech thru speech, presentation and repetition; and all others were of lesser activity. Likewise, all behaviours showed equal variety with the exception of didactic actions and presentation which were very lively. The performance indicators showed that the ratio of correct responses of individuals was moderately high; and that pronunciation, grammar and the combination of all three types of errors had the highest ratios (See Table 12).

TABLE 12

Lesson A: Activity and Variety Indicators

	ACTIVITY			VARIETY		
DIDACTIC BEHAVIOURS	SUM OF FREQUENCIES	TIME (in minutes)	ACTIVITY INDICATOR	SUM OF CATEGORIES	TIME (in minutes)	VARIETY INDICATOR
1. General Behaviour	472	25.30	18.6561	7	25.30	0.2767
2. Articulation	148	4.80	31.0000	5	4.80	1.0417
3. Presentation	17	0.41	41.5000	3	0.41	7.3171
4. Repetition	146	3.56	41.0112	4	3.56	1.1236
5. Verification	150	4.66	32.2000	7	4.66	1.5020
6. Didactic Actions	7	0.44	16.0000	4	0.44	9.0910
7. Speech thru Speech	47	0.89	53.0000	1	0.89	1.1236
8. Action and Speech	178	8.00	22.5316	7	8.00	0.8861

2.2 *Lesson B*

The two distinctions made of action and speech activities in Lesson B were classified as Didactic Behaviour and Language Behaviour. The didactic behaviours consisted of teaching and learning processes, and the language behaviours were concerned with the manner in which language was used. Whereas the didactic behaviours were divided into major categories and sub-categories, language behaviours distinguished only two categories: language use and language performance. There were two major categories and two sub-categories of teaching, whereas, there was only one category of learning.

The results show that the new didactic points had been used 662 times, and the total duration of these didactic activities was 25.3 minutes. And teaching behaviours occurred almost three times more often than learning behaviours, and the relative duration of teaching activities almost doubled that of learning. It was found that articulation, verification and action and speech had the highest frequency and duration, and although repetition was less frequent, it had been comparatively long. But speech thru speech was brief and infrequent. The relative duration and frequency of these behaviours were used to plot Behaviour Distribution Profiles of Lesson B (See Figures 19 to 28). In addition, the behaviour indicators were calculated from the data of the first analysis. And whereas Lesson B was very active and intense, it showed less variety. Learners' activities had been very active as well as varied; and speech thru speech, repetition and verification had been active but not varied. The performance indicators show that the ratio of correct individual utterance was moderately high; and the ratio of errors in pronunciation, vocabulary and grammar was high. Likewise, the ratio for vocabulary and grammar errors was high in Lesson B (See Table 13).

TABLE 13

Lesson B: Activity and Variety Indicators

	ACTIVITY			VARIETY		
DIDACTIC BEHAVIOURS	SUM OF FREQUENCIES	TIME (in minutes)	ACTIVITY INDICATOR	SUM OF FREQUENCIES	TIME (in minutes)	VARIETY INDICATOR
1. General Behaviour	662	25.27	26.1971	7	25.27	0.2770
2. Articulation	234	5.60	42.0000	5	5.60	0.8929
4. Repetition	174	3.50	50.0000	4	3.50	1.1429
5. Verification	260	5.28	49.2424	7	5.28	1.3258
7. Speech through Speech	14	0.24	58.3333	1	0.24	4.1667
8. Action and Speech	254	5.00	51.0000	7	5.00	1.4000

2.3 *Lesson C*

The action and speech activities of Lesson C had been analysed as Didactic and Language Behaviours. Didactic Behaviours measured teaching/learning processes, in terms of duration and frequency; whereas, Language Behaviours measured the quality of speech used by teacher and learner. In addition, didactic behaviours distinguished eight categories, and language behaviours, two.

The results showed that the new didactic points had been used 365 times, for the duration of 19.7 minutes. And teaching behaviours were twice as long, and occurred twice as often as learning behaviours. Moreover, repetition, action and speech, verification and articulation had been frequent and of long duration. Behaviour Distribution Profiles have been plotted from the results of the first analysis (See Figures 19 to 28). The data were used to calculate behaviour indicators of activity, variety, intensity and performance. The results show that Lesson C had been varied but not active, nor intense. And the performance indicators show that the proportion of learner's correct and incorrect utterances were high. Moreover, the analysis showed that didactic actions, repetition, speech thru speech and action and speech had been active and varied, and while presentation had been varied, it was not active (See Table 14).

TABLE 14

Lesson C: Activity and Variety Indicators

	ACTIVITY			VARIETY		
DIDACTIC BEHAVIOURS	SUM OF CATEGORIES	TIME (in minutes)	VARIETY INDICATOR	SUM OF FREQUENCIES	TIME (in minutes)	ACTIVITY INDICATOR
1. General Behaviour	7	20,00	0,3500	365	20,00	18,2500
2. Articulation	3	2,89	1,0381	97	2,89	34,0000
3. Presentation	6	1,11	5,4054	11	1,11	10,0000
4. Repetition	5	2,79	1,7921	174	2,79	62,4000
5. Verification	6	2,42	2,4793	120	2,42	50,0000
6. Didactic Actions	2	0,798	2,5063	16	0,80	20,0501
7. Speech thru Speech	2	0,25	8,0000	17	0,25	68,0000
8. Action and Speech	4	2,61	1,5300	123	2,61	47,1300

2.4 *Lesson D*

There were two distinctions made for action and speech activities in Lesson D: these were didactic and language behaviours. Whereas, didactic behaviours consisted in the analysis of the teaching/learning processes, language behaviours considered the manner in which language was used by teacher and learner during the lesson. The analysis of the didactic behaviours shows that there were three major categories and three sub-categories of teaching, whereas, there was only one category identified as learning.

And it was found that the new didactic points had been used 564 times, and, that they had been used for the duration of 37.1 minutes. The behaviours which were found to have the longest duration and frequency were: action and speech, articulation, presentation and didactic actions. But, the frequency and duration had been less for the remaining behaviours. The results of this first analysis were plotted onto Behaviour Distribution Profiles; and they were also used in the calculation of behaviour indicators of activity, variety, intensity and performance (See Figures 19 to 28). And whereas the low indicators show that Lesson D was not active, varied or intense, on the other hand, presentation had been very active, and articulation, repetition, verification and speech thru speech appeared to be highly varied. And the performance indicators show that the proportion of learner's correct utterances was low, but, that the proportion of incorrect utterances was high (See Table 15). Therefore, learners in Lesson D did not perform well during the lesson. This will be seen after comparing at the end of the analyses of the five lessons.

TABLE 15

Lesson D: Activity and Variety Indicators

	ACTIVITY			VARIETY		
DIDACTIC BEHAVIOURS	SUM OF FREQUENCIES	TIME (in minutes)	ACTIVITY INDICATOR	SUM OF CATEGORIES	TIME (in minutes)	VARIETY INDICATOR
1. General Behaviour	564	37.11	15.1981	9	37.11	0,2425
2. Articulation	154	5,02	31.0000	6	5,02	1.1952
3. Presentation	184	5.10	36.1000	8	5.10	1.5686
4. Repetition	85	2,13	40.0000	6	2,13	2,8169
5. Verification	156	3,79	41.1609	6	3.79	1.5831
6. Didactic Actions	95	8,69	11,0000	8	8,69	0,9206
7. Speech thru Speech	2	0,04	50,0000	1	0,04	25,0000
8. Action and Speech	183	10,12	18,0741	6	10,13	0,5923

2.5 *Lesson E*

The action and speech activities of Lesson E had been divided into Didactic and Language Behaviours. Whereas, didactic behaviours analysed the teaching/learning processes, language behaviours considered the manner in which language was used by the teacher and learner. The results show that there were three major categories and three sub-categories identified for teaching, but, there was only one category identified as learning.

The analysis showed that the new didactic points had been used 567 times and the total duration of these behaviours was 32.12 minutes. The most frequently used behaviours were articulation, verification and action and speech; these had been also of long duration. And although the remaining behaviours had been brief and short, repetition had been comparatively frequent. The results of this first analysis have been plotted onto Behaviour Distribution Profiles; they were also used to calculate behaviour indicators of activity, variety, intensity and performance (See Figures 19 to 28). Lesson E was neither active nor intense, but it had been quite varied. However, articulation, presentation, repetition, verification and speech thru speech had been very active and varied, but whereas action and speech had been varied, it lacked activity. And, verification and speech thru speech were active but not varied. And, the performance indicators showed that the ratio of correct as well as incorrect utterances was high, buth that the ratio of errors in vocabulary and grammar errors was low (See Table 16).

TABLE 16

Lesson E: Activity and Variety Indicators

	ACTIVITY			VARIETY		
DIDACTIC BEHAVIOURS	SUM OF FREQUENCIES	TIME (in minutes)	ACTIVITY INDICATOR	SUM OF CATEGORIES	TIME (in minutes)	VARIETY INDICATOR
1. General Behaviour	567	32.13	18.0000	10	23.13	0.3112
2. Articulation	200	4.44	45.0450	5	4.44	1.1261
3. Presentation	56	1.77	32.0000	4	1.77	2.2600
4. Repetition	118	1.62	73.0000	3	1.62	1.8520
5. Verification	226	4.20	54.0000	6	4.20	1.4300
6. Didactic Actions	30	2.94	10.2041	4	2.94	1.4000
7. Speech thru Speech	44	0.56	79.0000	1	0.56	1.8000
8. Action and Speech	226	7.01	32.2400	8	7.01	1.1412

The presentation and interpretation of the results of the polychronometric analyses of the five ESL classrooms having been completed, we can now compare the results of the analyses of these five lessons. First let us compare their profiles.

3. Behaviour Distribution Profiles

How to read the profile

Horizontal axis gives time in minutes and seconds.
Vertical axis gives frequency of components.
(Each line records one occurence)

▤ Occurences in the second language (the language being taught).

▥ Each broken line represents ten occurences in the second language.

■ Occurences in the language of the learner (the mother tongue).

3.1 *Didactic Behaviour* (Overview)

This profile gives the overall distribution of action and speech activities between the teacher and the class - teaching activities and learning activities. They include the following.

Teaching

AR	(Articulation includes linking with other lessons)
SP	(Semantic Presentation, including new meanings)
FP	(Formal Presentation, teaching pronunciation)
FR	(Formal Repetition, as in songs)
SR	(Semantic Repetition, as in games, drills)

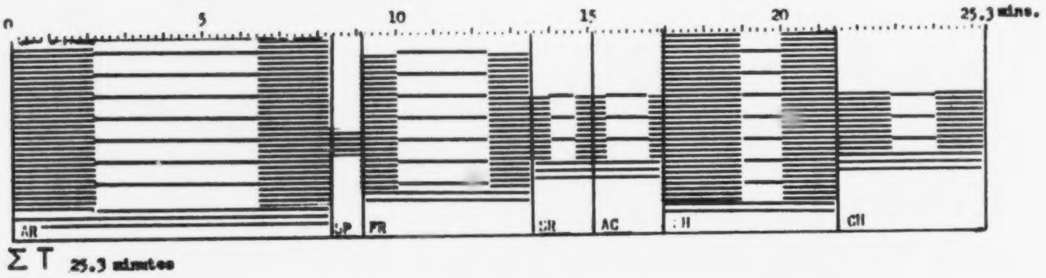
Learning

AC	(Action responses of learner)
SH	(Speech responses of learner)
SM	(Speech accompanied by significant motion)
RE	(Recitation or reading aloud)
CH	Choral or group speech or responses)

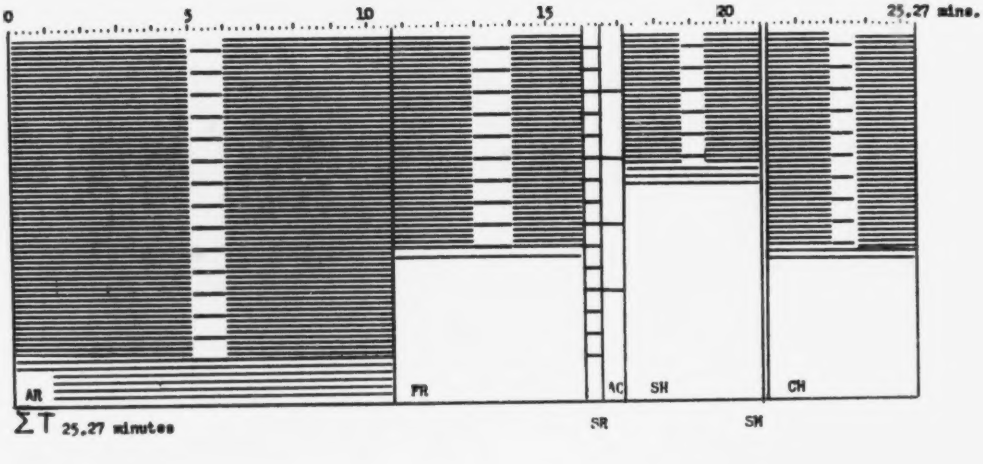
FIGURE 19

Distribution Profile of Didactic Behaviours (Overview)

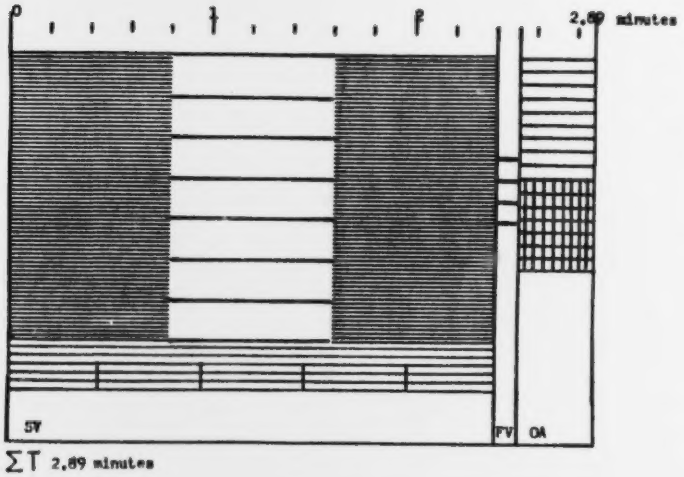
Lesson A



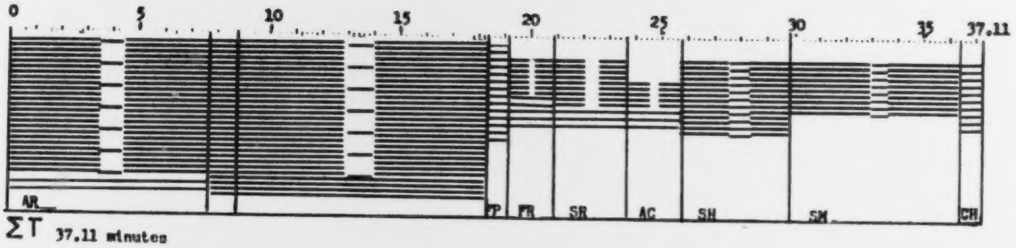
Lesson B



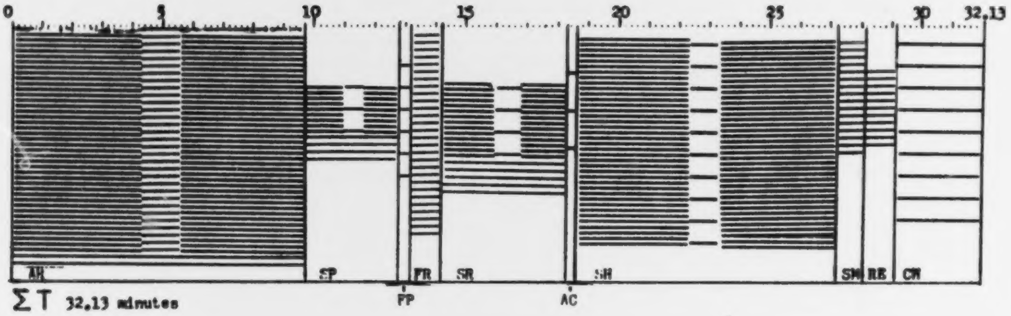
Lesson C



Lesson D



Lesson E



3.2 *Articulation*

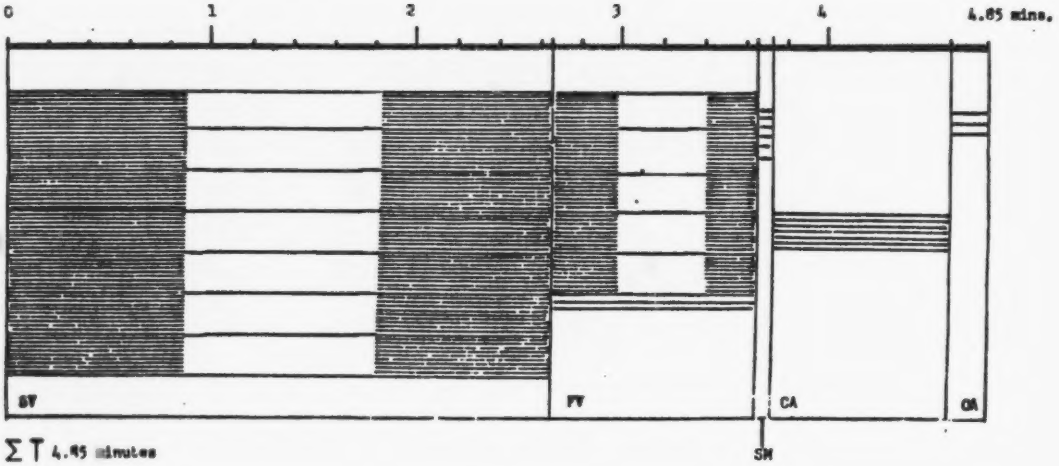
Articulation refers to the manner in which the teacher checks learners progress as well as the method of linking known materials with the new ones being taught. The different types of articulation are:

SV	(Semantic Verification: checking for understanding)
FV	(Formal Verification: checking pronunciation)
WV	(Written Verification: checking writing & spelling)
AM	(Active Motivation: attention thru actions)
SM	(Speech Motivation: talk to attract attention)
LK	(Linking known to unknown)
DR	(Dramatization: sketches & skits)
GC	(Guided conversation and dialogues)
CA	(Class Administration)
OA	(Other activities not identified in this segment)

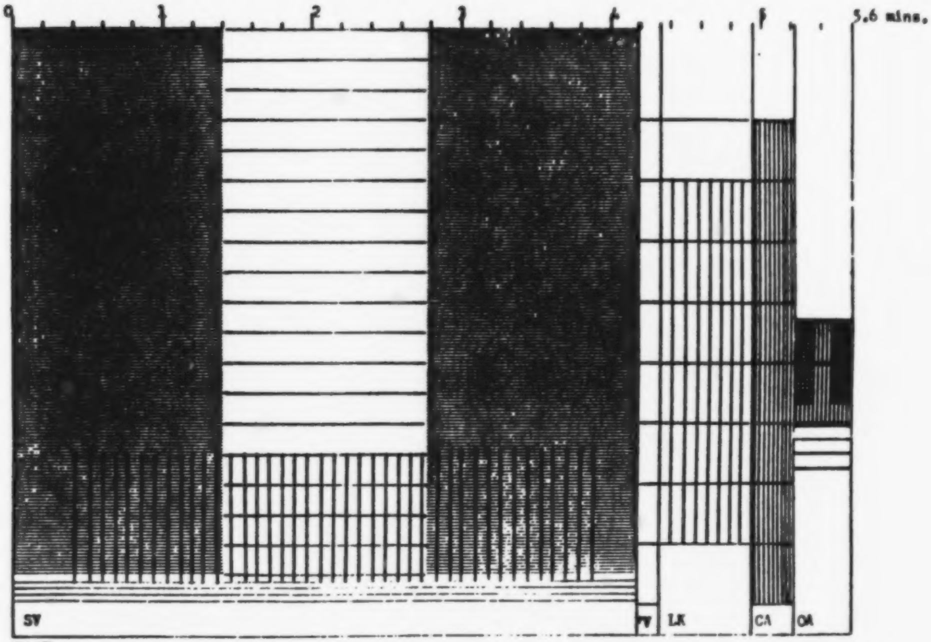
FIGURE 20

Distribution Profiles of Articulation

Lesson A

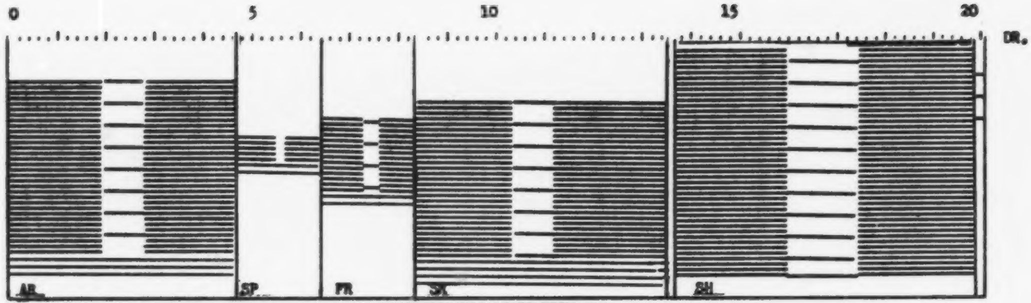


Lesson B



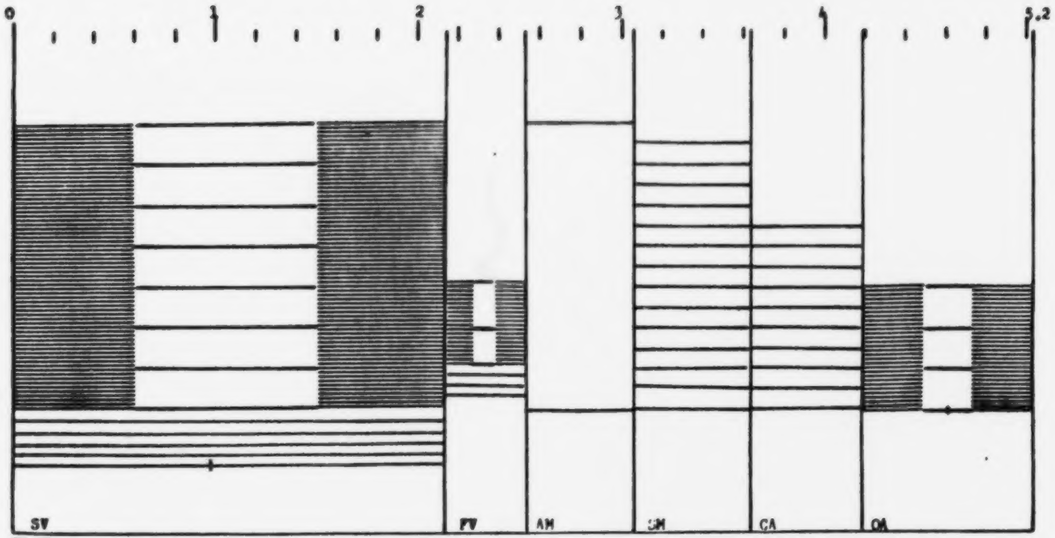
$\sum T$ 5.60 minutes

Lesson C



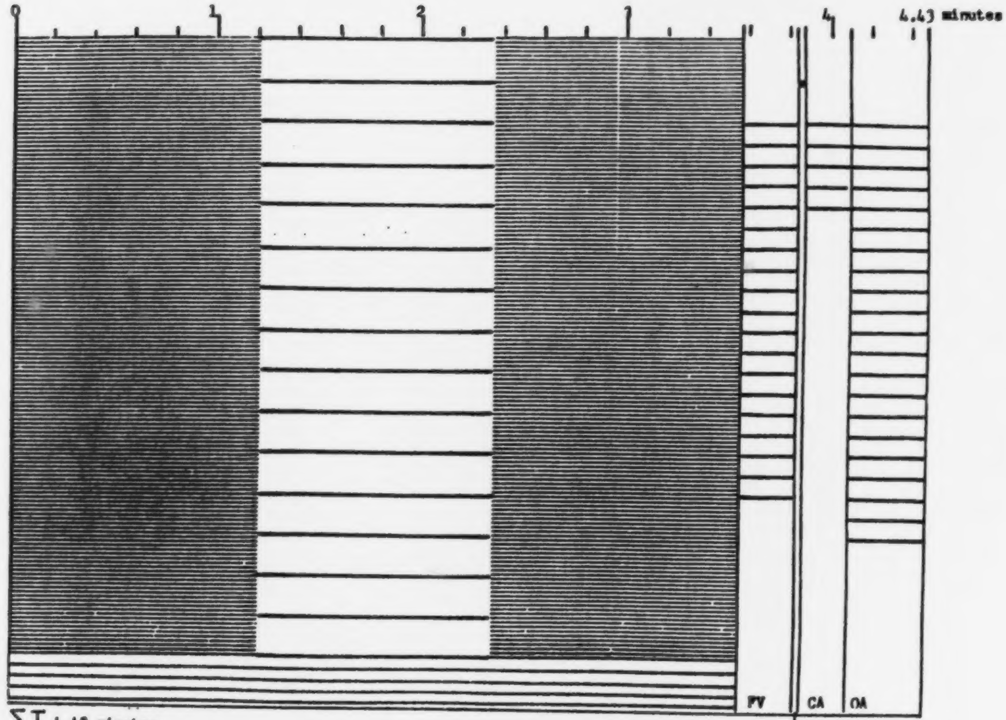
$\sum T$ 20.0 minutes

Lesson D



ΣT 5.02 minutes

Lesson E



ΣT 4.43 minutes

3.3 *Presentation*

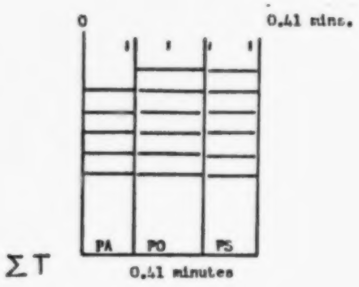
Presentation. The manner in which the new didactic points are introduced to learners for the first time. This may be (a) Ostensive or meaning conveyed through objects, e.g. toys, books, things. (b) Pictoral, teaching meaning thru wall charts, flash cards and films. (c) Contextual or the use of translation e.g. fille, girl. C'est une fille. It's a girl.

PA	(Presentation thru actions and speech)
PO	(Presentation thru objects and speech)
PS	(Presentation thru L2 speech only)
PX	(Presentation thru pictures)
PF	(Presentation thru films)
PT	(Presentation thru translation)
EX	(Expaining new forms and rules)
OM	(Oral modelling)
WE	(Writing on chalk board)
OP	(Other activities during Presentation)

FIGURE 21

Distribution Profiles of Presentation

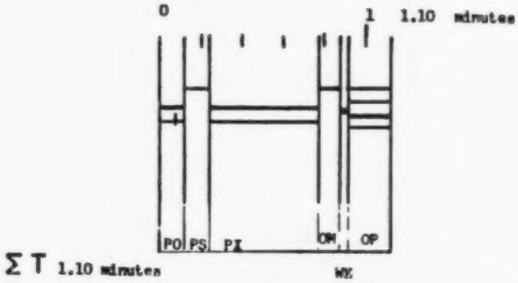
Lesson A



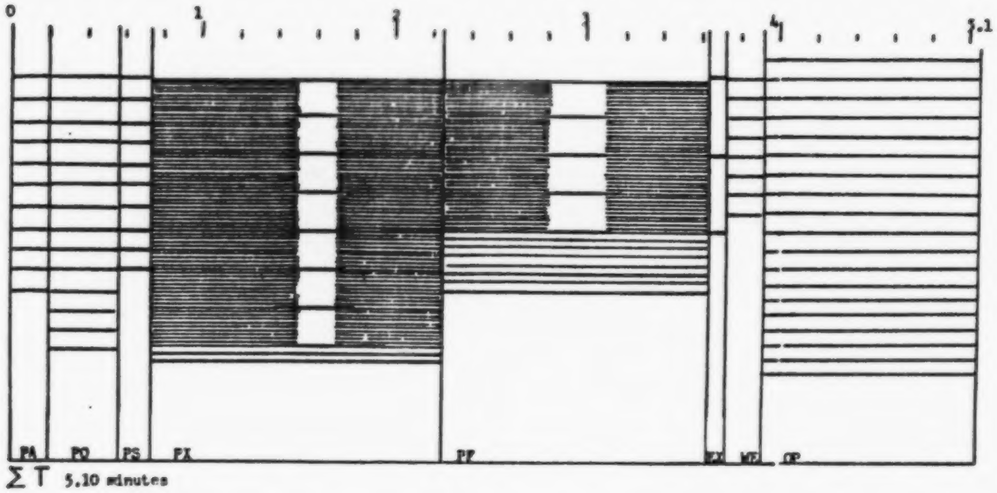
Lesson B

NONE

Lesson C

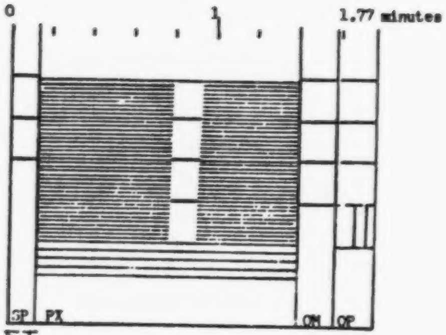


Lesson D



ΣT 5.10 minutes

Lesson E



ΣT 1.77 minutes

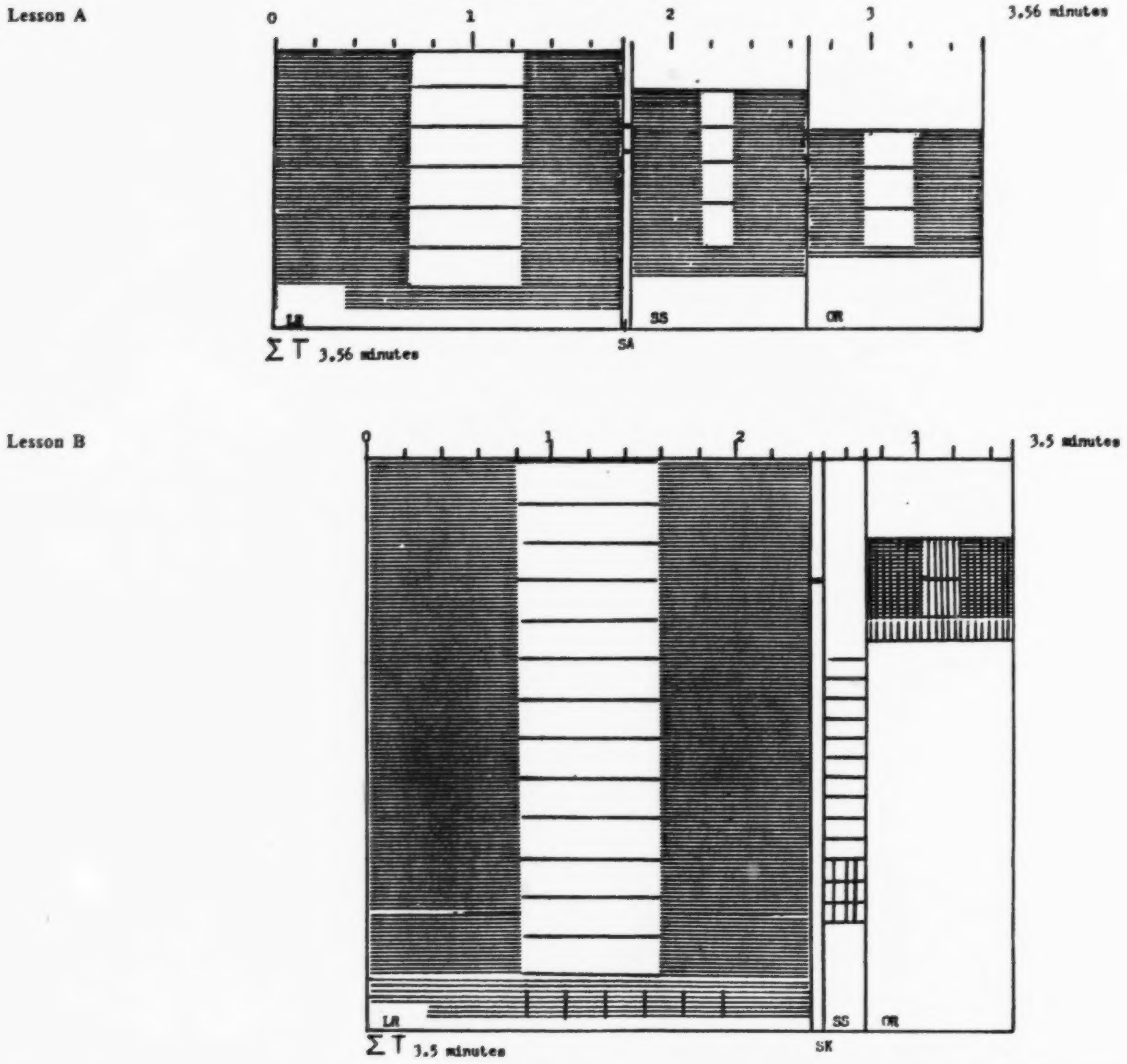
3.4 *Repetition*

Repetition. This is the technique used by the teacher in order to reinforce the new elements of vocabulary and grammar. The types used in the study are:

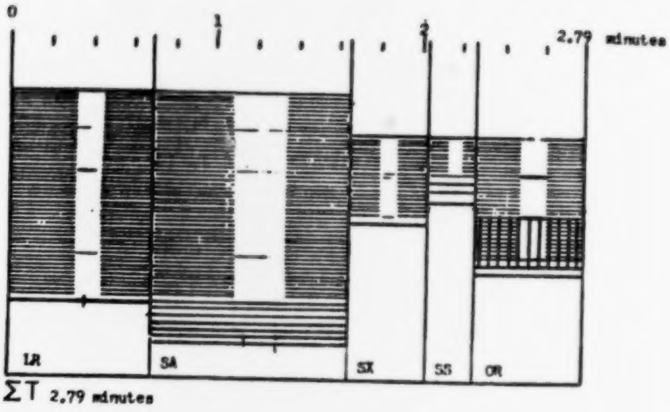
LR	(Listen-and-repeat, e.g. imitation drills)
LW	(Listen-and-write, e.g. dictation)
KW	(Look-and-write, e.g. copying)
RS	(Read-and-say, e.g. reading aloud)
CR	(Choral repetition)
ST	(Speech thru translation)
SA	(Speech thru actions & objects)
SX	(Speech thru pictures)
SS	(Speech thru Speech, drills)
OR	(Other activities in Repetition)

FIGURE 22

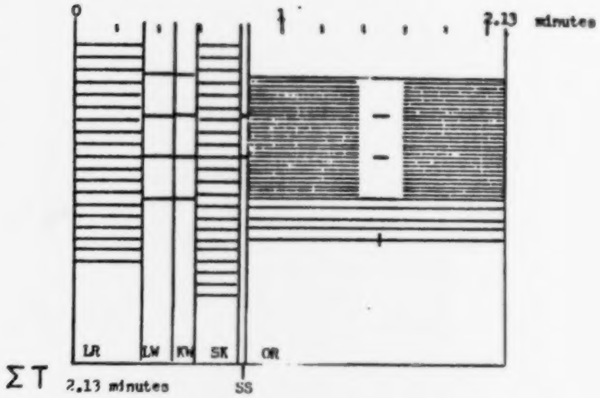
Distribution Profiles of Repetition



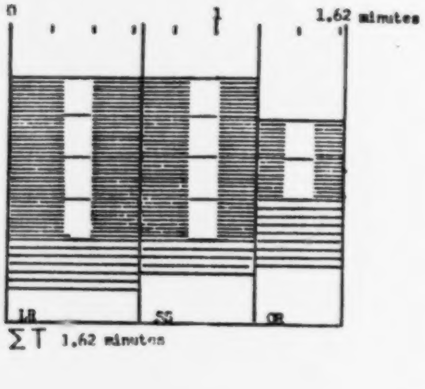
Lesson C



Lesson D



Lesson E



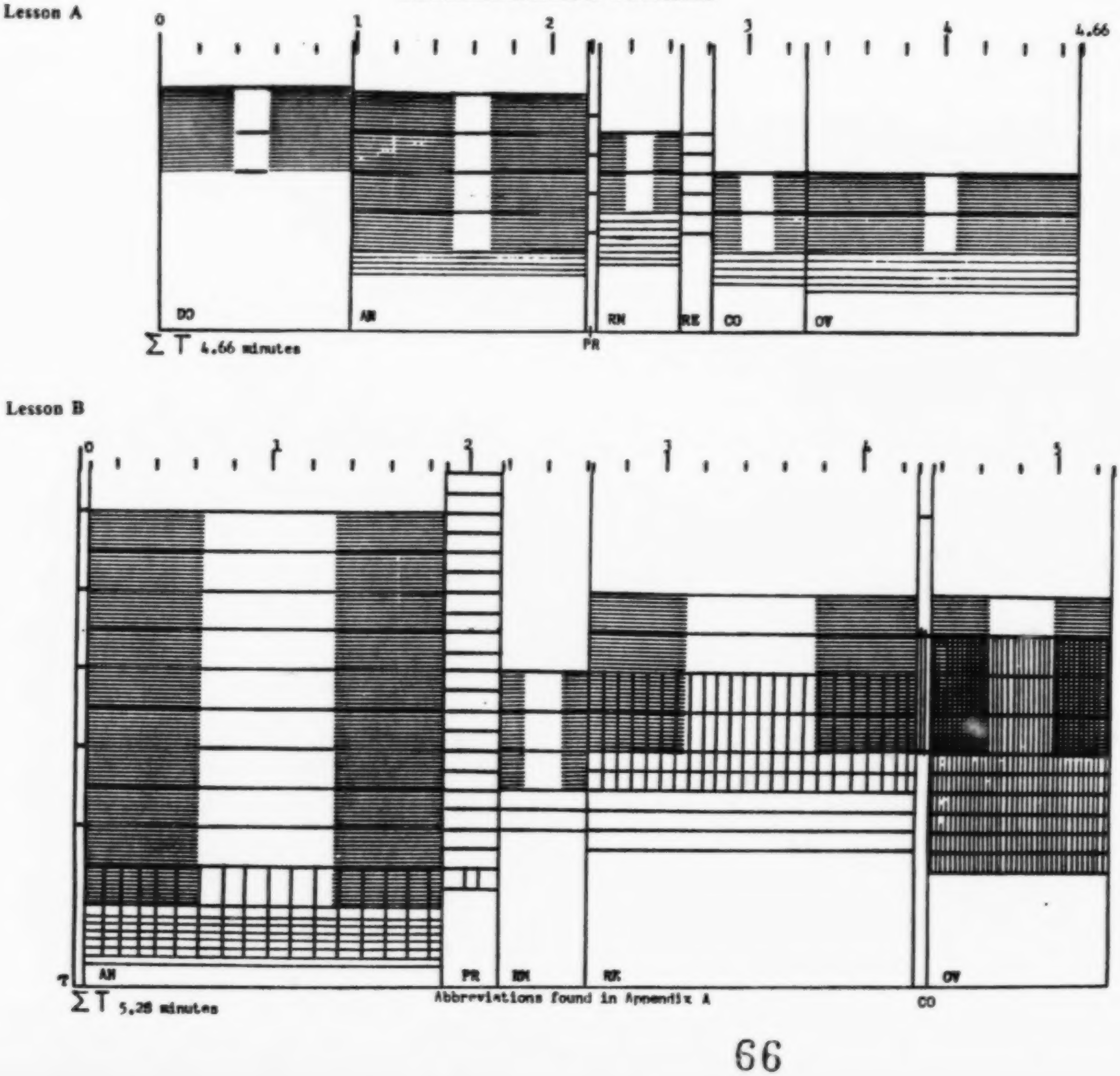
3.5 *Verification*

Verification. This is the technique used by the teacher in order to check pupils' understanding of the new didactic points and to re-explain points which are not fully grasped by them. The techniques identified in this study are:

- PT (Point: checking by pointing)
- DO (Do: checking through actions)
- AN (Answers: checking thru questions)
- WR (Write: checking thru writing)
- PR (Prompting by teacher)
- RM (Rephrasing & remodelling by teacher)
- RE (Re-explaining orally new points)
- WB (Writing on chalk board)
- CO (Confirming responses as acceptable)
- OV (Other activities during Verification)

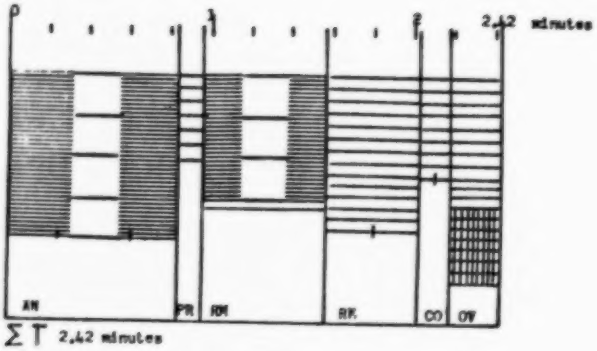
FIGURE 23

Distribution Profiles of Verification

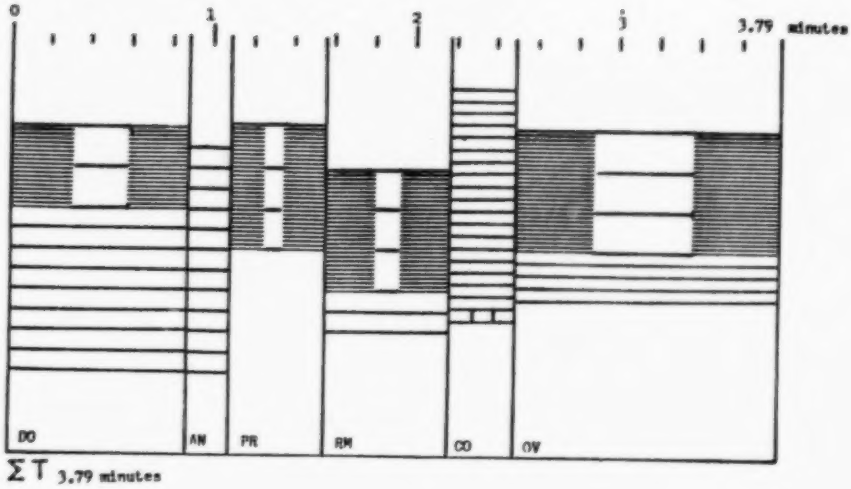


Abbreviations found in Appendix A

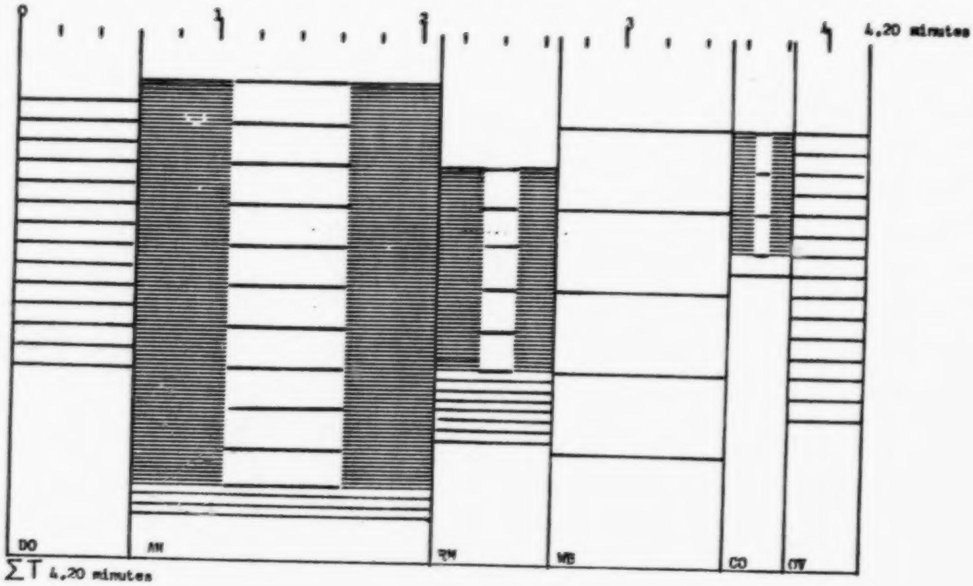
Lesson C



Lesson D



Lesson E



3.6 *Didactic Action*

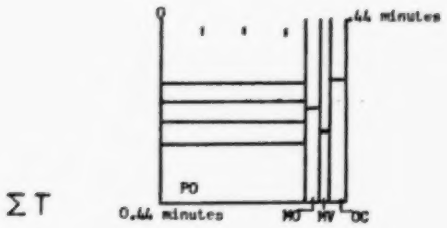
Actions (didactic). These are actions used by the teacher during the presentation of the new didactic points. The purpose of these actions are to enable pupils to grasp the meaning of the new items and to learn the pronunciation thru motions of the hand. The didactic actions identified in the study are:

WB	(Writing on chalk board)
PE	(Pointing to board)
PX	(Pointing to pictures)
PO	(Pointing to objects)
MO	(Manipulating objects)
MV	(Movement, i.e. going to board, to desk, or any meaningful movement)
MI	(Miming)
IA	(Interacting non-verbally with pupil)
MM	(Manipulating Machine, e.g. overhead projector recorder on so on)
OC	(Other didactic actions during Presentation)

FIGURE 24

Distribution Profiles of Actions (Didactic)

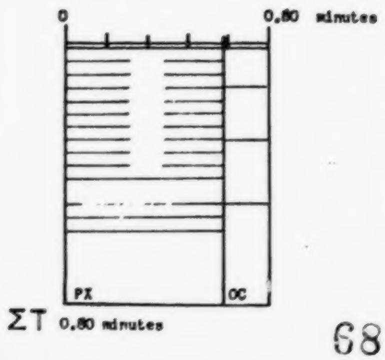
Lesson A



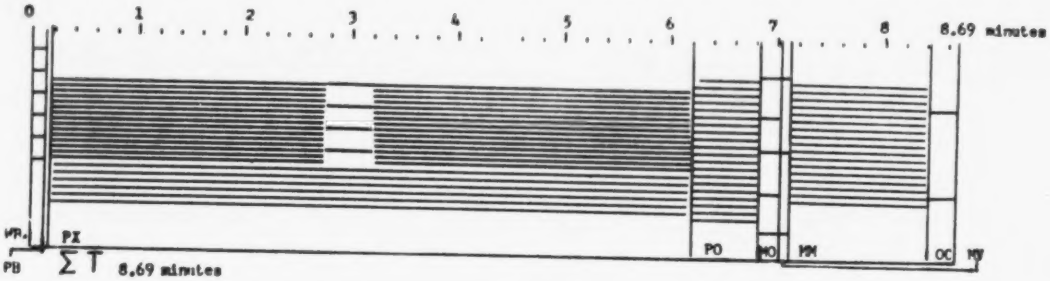
Lesson B

NONE

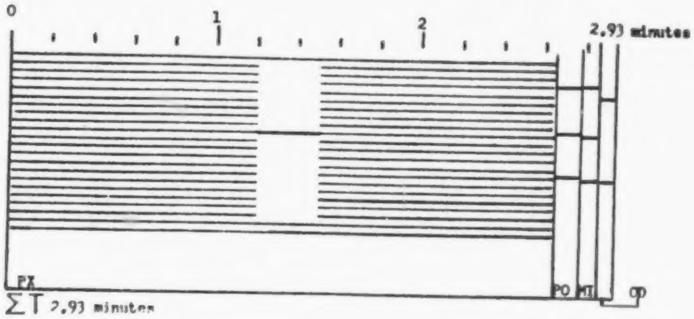
Lesson C



Lesson D



Lesson E



3.7 *Speech thru Speech*

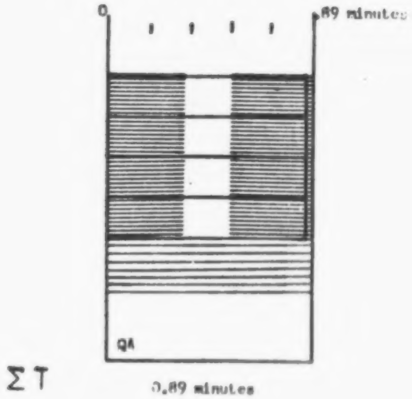
Speech thru Speech. This has been identified as drills in the form of question-and-answers and substitution. The following are the types used:

PP	(Pattern practice, e.g. substitution drills)
QA	(Question-and-answer drills)
GG	(Guessing games)
OG	(Oral composition Games)
OD	(Oral composition drills)
RD	(Reproduction drills)
DE	(Description and exposition exercises)
DM	(Dialogue modification)
TT	(Topical talks and prepared speaches)
OS	(Other activities occurring during this set)

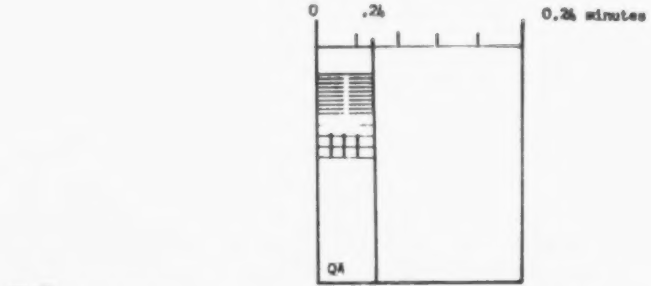
FIGURE 25

Distribution Profiles of Speech thru Speech

Lesson A

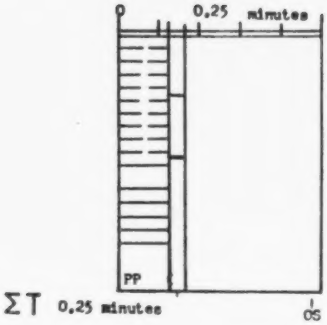


Lesson B

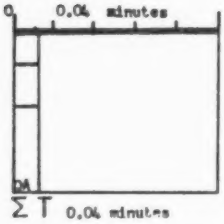


Abbreviations found in Appendix A

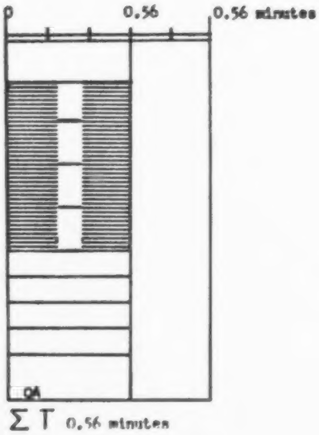
Lesson C



Lesson D



Lesson E



3.8 *Action and Speech*

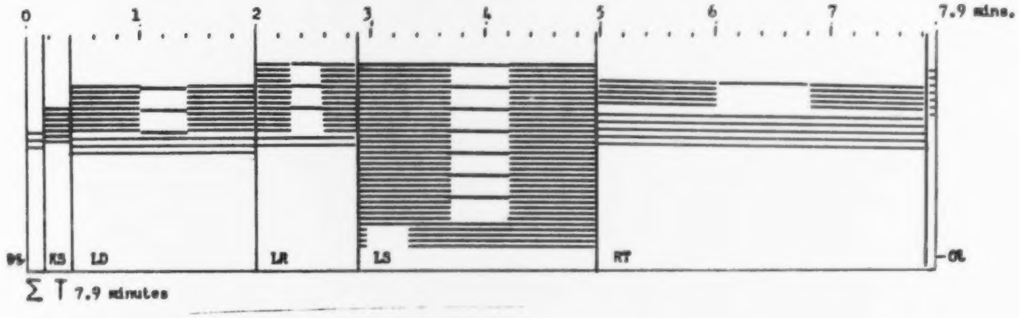
Action and Speech. This refers to learners' individual and group didactic activities during the lesson. The types of action and speech identified are as follows:

DS	(Do and say, identifying something and naming it)
KS	(Look and say, looking at something and naming it)
LD	Listen and do, action response to teacher's command)
LR	(Listen and repeat, verbatim repetition)
LS	(Listen and say, learners' answers to questions)
LW	(Listen and write, dictations or drawing)
TR	(Translated)
RC	(Recite)
RT	(Repeat together)
OL	(Other activities during Action and Speech)

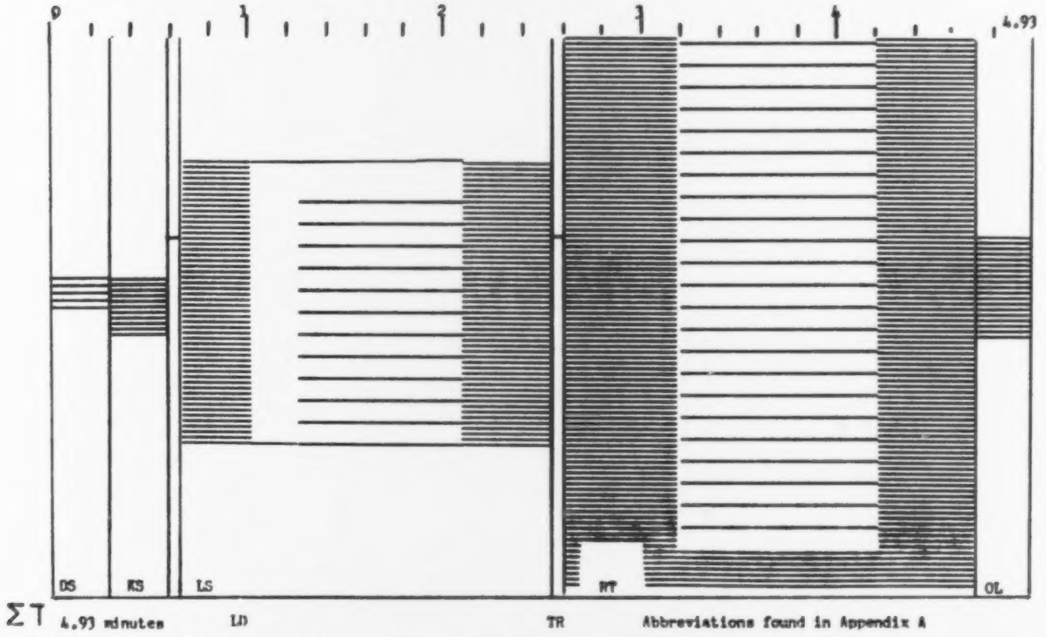
FIGURE 26

Distribution Profile of Action and Speech

Lesson A

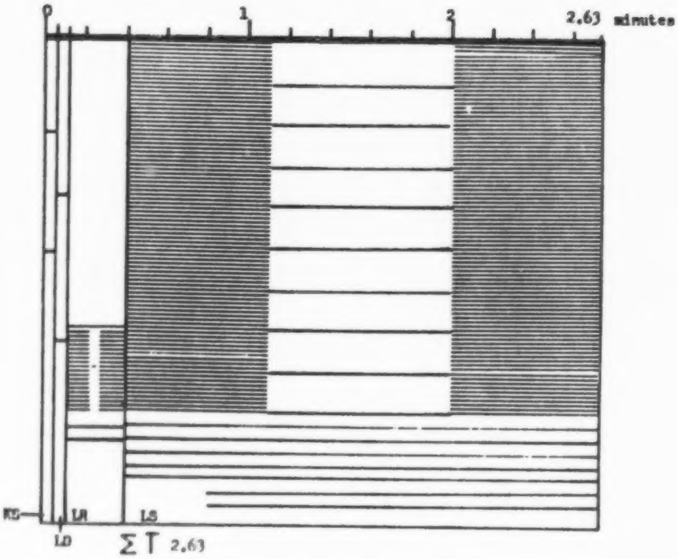


Lesson B

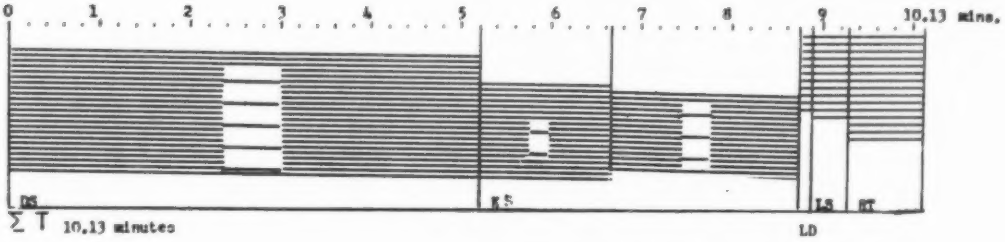


Abbreviations found in Appendix A

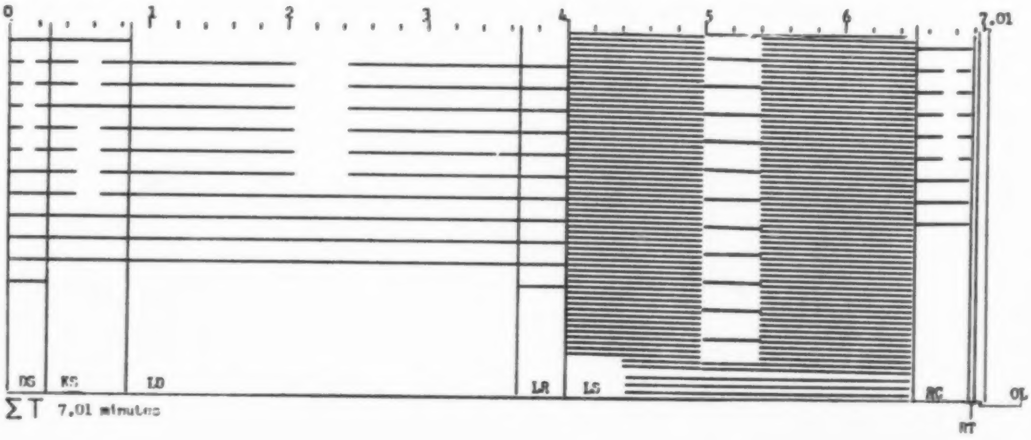
Lesson C



Lesson D



Lesson E



3.9 *Language Use*

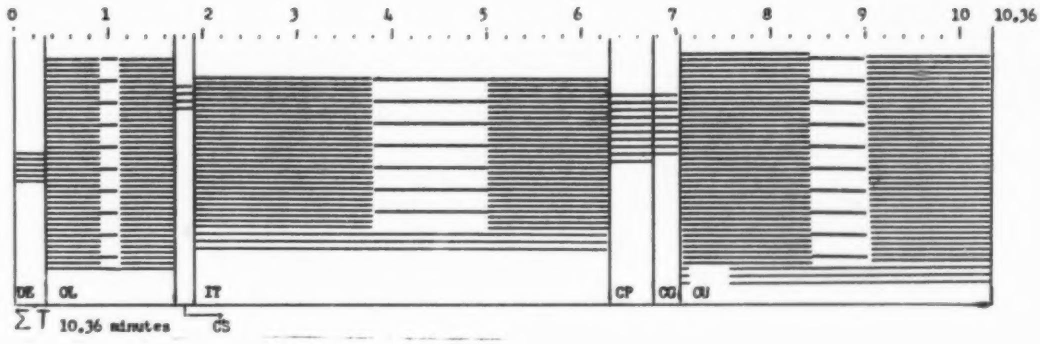
Language Use. This behaviour has been identified in terms of the teachers' use of language during the lesson. The techniques given below are self explanatory. They are as follows:

DE (Demonstration)
ER (Explanation of rules of lecturing in L2)
MT (Using the Mother tongue)
GD (Giving Dictation)
QL (Questioning a learner)
CS (Completing a learner's statement)
IT (Iteration, as in repeating pupils' responses)
CP (Correction of Pronunciation)
CG (Correction of Grammar and/or vocabulary)
OU (Other uses of language)

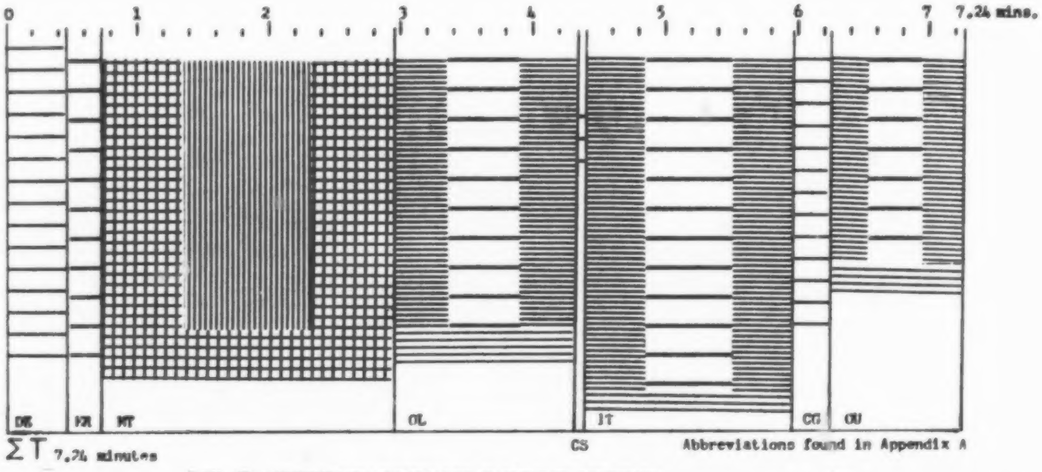
FIGURE 27

Distribution Profiles of Language Use

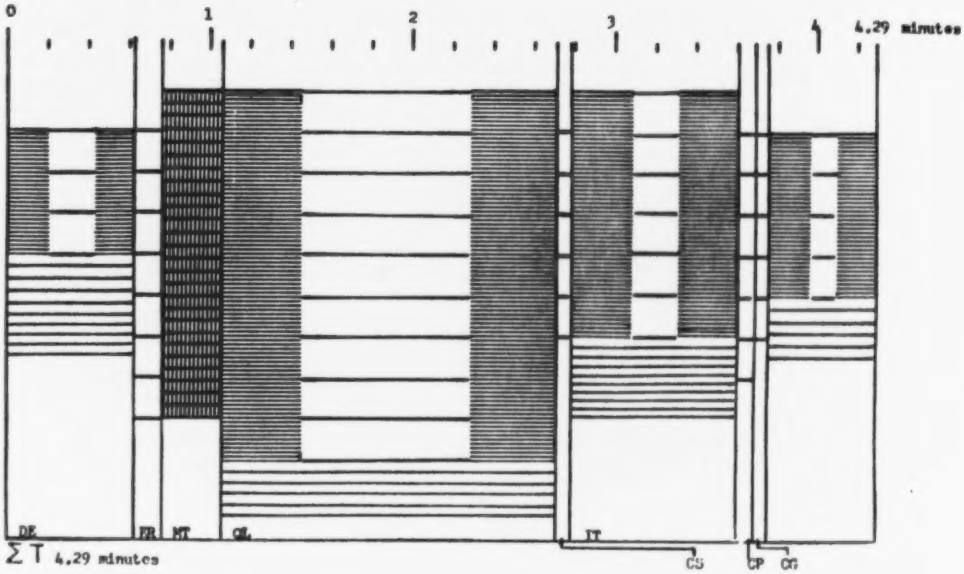
Lesson A



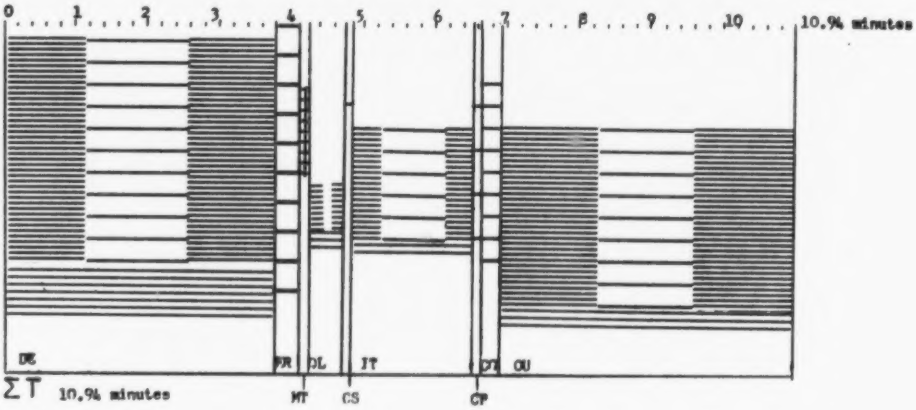
Lesson B



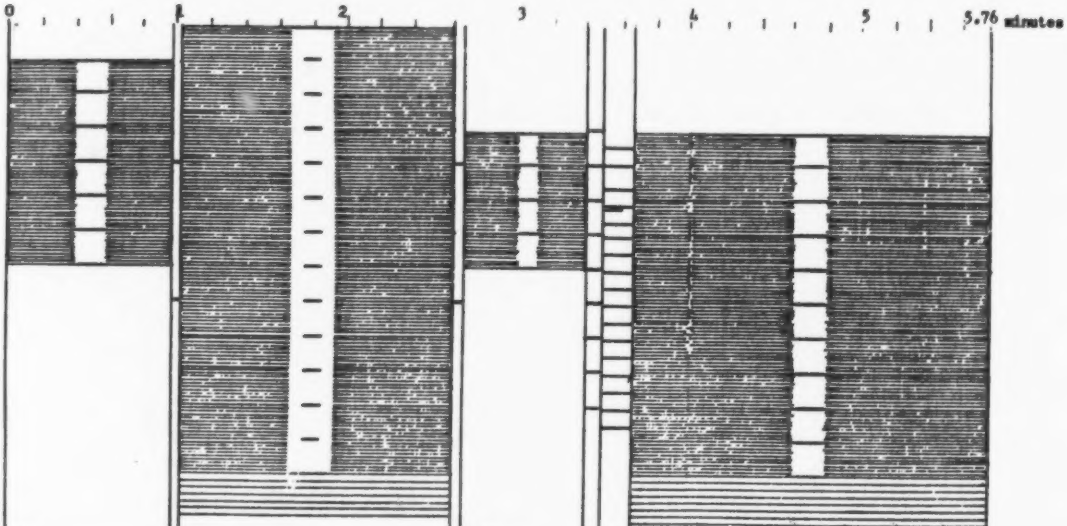
Lesson C



Lesson D



Lesson E



3.10 *Language Performance*

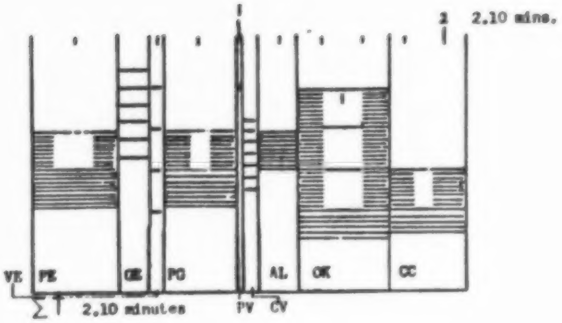
Language Performance (individual). This set includes all individual responses of learners, both correct and incorrect. The individual responses identified in this study are as follows:

PE	(One or more phonetic error)
GE	(One or more grammar error)
VE	(One or more vocabulary error)
PG	(Pronunciation & grammar error)
PV	(Pronunciation and vocabulary error)
GV	(Grammar & vocabulary error)
AL	(All three types of error)
OK	(Everything correct in a dependent utterance)
CC	(Correct response after teacher's correction)
SI	(Statement independently generated and error-free)

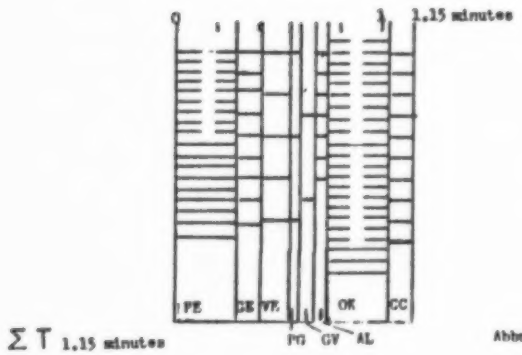
FIGURE 28

Distribution Profiles of Language Performance

Lesson A

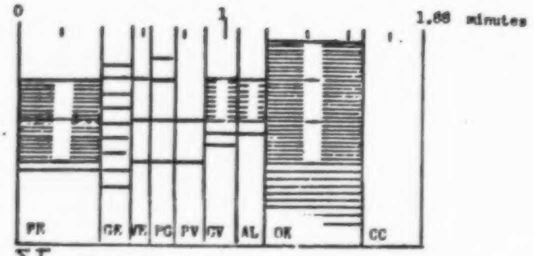


Lesson B



Abbreviations found in Appendix A

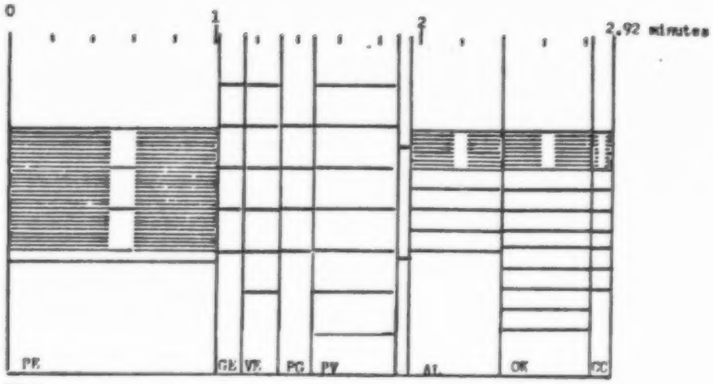
Lesson C



ΣT 1.86 minutes

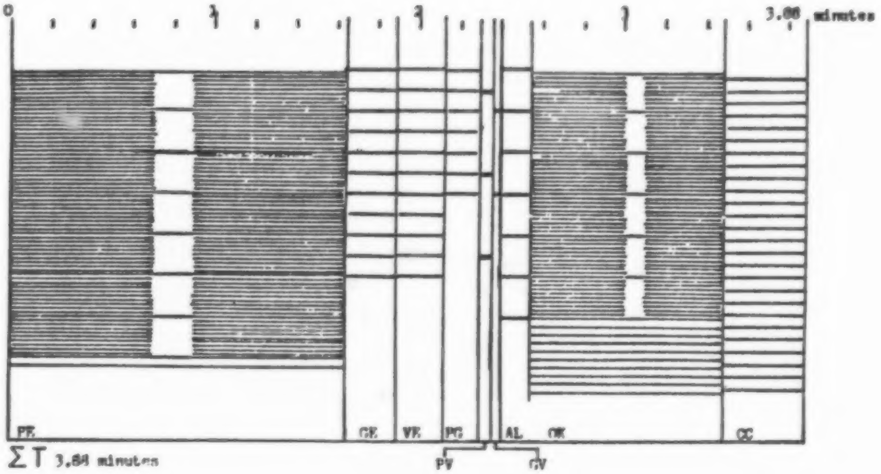
Abbreviations found in Appendix A

Lesson D



ΣT 2.92 minutes

Lesson E



ΣT 3.88 minutes

Chapter 4

LEARNING AND TEACHING BEHAVIOUR: CORRELATING LESSON COMPONENTS

1. **Comparative Component Analysis**
    1.1 *Duration of Frequency of Components*
    1.2 *Indicators of Teaching Behaviour*

2. **Relating Teaching to Learning**
    2.1 *Overall Classroom Behaviour*
    2.2 *Didactic Behaviours*
    2.3 *Language Use*
    2.4 *Survey Summary of Correlations*

3. **Interpretation of Teaching and Learning Differences**
    3.1 *Correlation of Behaviour and Mean Gain Scores*
    3.2 *Findings and Conclusions*

1. Comparative Component Analysis

The five lessons can now be compared, in terms of relative duration and frequency of didactic and language behaviours, and then according to behaviour indicators.

1.1 *Duration of Frequency of Components*

The polychronometric analyses of the five lessons have already been presented separately in the preceding paragraphs. In this section the writer will attempt to compare and contrast the relative duration and frequency of components found for Lessons A, B, C, D and E.

Table 22 shows the relative duration of the five lessons. The analyses show that the duration of Lessons A and B was equal; the new didactic points were taught for the duration of 25.3 minutes. Whereas, Lesson D had been of the longest duration, 37.1 minutes, Lesson C was conducted in half this time; it lasted 19.7 minutes, and the duration of Lesson E was 32.12 minutes, making this lesson second to that of Lesson D.

However, the new didactic points had been used with the highest frequency in Lesson B; they occurred 662 times as compared with Lesson E with 567 occurrences, Lesson D, 564, Lesson A, 472 and least of all in Lesson C with 365 occurrences.

And, although Lessons A and B had been of equal length, the new materials were the most frequently used in B than in A; and also less frequent in the three other lessons. However, Lesson D had been of the longest duration. Lesson C had been the shortest, and Lesson E was the second in length.

Table 22 also shows the relative duration and frequency of the teaching and learning behaviours. And the relative duration of teaching for each of the five lessons, was as follows: A, 15.12 minutes, B, 16.4 minutes, C, 13.4 minutes, D, 23.60 minutes and E, 18.5 minutes.

TABLE 17

The Relative Duration and Frequency of Components in Five Lessons

LESSONS	NUMBER	TOTAL DURATION (in minutes)			BEHAVIOURS FREQUENCY		
		LESSON (Dr.)	TEACHING (Dr.)	LEARNING (Dr.)	LESSON (Fr.)	TEACHING (Fr.)	LEARNING (Fr.)
A	23	25.30	15.12	10.14	472	274	198
B	24	25.30	16.40	8.90	662	407	255
C	21	19.70	13.40	6.32	365	240	125
D	26	37.10	23.60	13.50	564	403	161
E	25	32.12	18.45	13.67	567	347	220

Dr. duration
Fr. frequency

The relative duration of learning for each of the lessons: A, 10.14 minutes, B, 8.90 minutes, C, 6.32 minutes, D, 13.50 minutes and E, 13.67 minutes.

And whereas, teaching behaviours lasted twice as long as learning, in Lessons D, B and C, the duration of teaching and learning in Lessons A and E was equal. Figure 67 (a) and (b) illustrate clearly these similarities and differences.

Table 17 shows the frequency of teaching and learning behaviours. The frequency of teaching for the five lessons: A, 274, B, 407, C, 240, D, 403 and E, 347.

Frequency of learning for each of the five lessons: A, 198, B, 255, C, 125, D, 161 and E, 220.

The new didactic points were used with the highest frequency by teacher B, but they were the least frequently used by teachers C and A. And teachers D and E used them almost as frequently as teacher B.

There were similarities shown for learning: learners in Lessons B, E and A made more frequent use of the new materials than learners in Lessons C and D. But, whereas teacher behaviours were highly frequent in D but less in C, the behaviours of learners were the least frequent in these two lessons.

Table 18 and Figures 29 and 30 (a) and (b) show the duration and frequency of the specific behaviours. The presentation of these behaviours will follow the order on the Table and Figures.

Articulation. The relative duration for each of the five lessons: A, 4.85 minutes, B, 5.60 minutes, C, 2.89 minutes, D, 5.02 minutes and E, 4.44 minutes. The results show that articulation was used longer in Lesson B and the time spent in this activity was equal in Lessons A and D. But in Lesson C, it was the shortest. The duration in Lesson E was between these two extremes.

The frequency of articulation for each of the five lessons: A, 148, B, 234, C, 97, D, 154, E, 200. The new materials had been used very frequently during articulation, in Lessons B and E, but to a lesser degree during Lessons D, A and C. Therefore, articulation behaviours had been used more frequently and for a longer time in Lesson B than in the other lessons.

Presentation. The relative duration for each of the four lessons: A, 0.41 minutes, C, 1.10 minutes, D, 5.10 minutes, E, 1.77 minutes. This set had been identified in four classrooms. The results show that the new materials had been presented for a longer time, to learners in Lesson D; and whereas this time-slot was short for the other lessons, Lesson A had been the briefest. Moreover, vocabulary and grammar were taught by teacher D five times longer than in Lesson C; and compared with Lesson A, the duration of presentation had been 12 times longer in Lesson D.

The frequency of presentation for each of the four lessons: A, 17, C, 11, D, 184, E, 56. The new didactic points had been used with the highest frequency in Lesson D and less so in Lesson E, but the least occurrences of the new materials were during Lessons A and C. Therefore, the new materials had been used more often and for a longer time in Lesson D than in the other lessons.

Repetition. The relative duration for each of the five lessons: A, 3.56 minutes, B, 3.50 minutes, C, 2.79 minutes, D, 2.13 minutes, E, 1.62 minutes. The new didactic points had been reinforced for a longer time during Lesson A, B and C but were of shorter duration during Lessons D and E.

The frequency of repetition for each of the five lessons: A, 146, B, 174, C, 174, D, 85, E, 118. Repetition had been used with equal frequency during Lessons B and C, but less frequently in Lessons A and E, and the least often used in Lesson D. However, roles were changed between teacher and learner; therefore, this might account for the low occurrences by teacher D. And, whereas repetition during Lessons A and B had been of equal duration, the new materials were repeated with the highest frequency in Lessons B and C. Moreover, teacher B used repetition as a means of introducing the new vocabulary and grammar to learners.

Verification. The relative duration for each of the five lessons: A, 4.66 minutes, B, 5.28 minutes, C, 2.42 minutes, D, 3.79 minutes, E, 4.20 minutes. Teachers B and A checked learners' language acquisition for a longer time than teachers in the other classrooms. And, the duration of verification was also three times longer than in Lesson C.

The frequency of verification for each of the five lessons: A, 150, B, 260, C, 120, D, 156, E, 226. Teacher B checked learners' progress more often than teachers in the four other lessons. However, verification was frequently used by teacher E but to a lesser degree by teachers A, C and D. Therefore, verification had been of the highest frequency and duration in Lesson B. In addition, questioning learners occurred 109 times in Lesson B and 103 times in Lesson E, whereas, this behaviour was repeated less frequently in Lesson A; teacher A questioned learners 44 times, in Lesson D, 31 occurrences, and least of all, in Lesson C, where 12 questions had been asked. The other behaviours of verification had been infrequently used, with the exception of corrections of vocabulary and grammar. The teacher's rate of correcting had been highly frequent in Lessons B, C, D and E, but teacher A made the least number of corrections.

Actions (Didactic). The relative duration for each of the four lessons: A, 0.44 minutes, C, 0.80 minutes, D, 8.69 minutes, E, 2.94 minutes. The duration of didactic actions during presentation was longest in Lesson D. Teacher D used eight different types of actions, but the most dominant was pointing to pictures; this occurred 46 times, for the duration of 6.04 minutes; the other seven behaviours were of lesser duration. In Lesson C, there was only one action performed, and there were four didactic actions in Lessons A and E. The most dominant behaviour which was used by all teachers, was pointing to pictures.

The frequency of didactic actions for each of the four lessons: A, 7, C, 16, F, 95, E, 30. Didactic actions had been used with the highest frequency in Lesson D, and to a lesser degree in Lessons E, C and A; but they were the most infrequently used in Lesson A.

Speech thru Speech. The relative duration of speech thru speech for each of the five lessons: A, 0.89 minutes, B, 0.24 minutes, C, 0.23 minutes, D, 0.04 minutes, E, 0.56 minutes. The duration of drills in all five lessons had been short; however, it was comparatively longer in Lessons A and E, and whereas, speech thru speech lasted for the same length of time in Lessons B and C, it was very brief in Lesson D.

The frequency of speech thru speech for each of the five lessons: A, 47, B, 14, C, 17, D, 2, E, 44. The dominant behaviour used in this set, by all five teachers, was repetition of the new didactic points by question-and-answer technique. This technique was used very frequently by teacher A and to a lesser degree by teacher E, and while teachers B and C used it less frequently, teacher D rarely used this technique.

Action and Speech. The relative duration of action and speech for each of the five lessons: A, 7.90 minutes, B, 4.93 minutes, C, 2.61 minutes, D, 10.13 minutes, E, 7.01 minutes. Learners' activities (action and speech) had been the longest in Lesson D, and to a lesser degree in Lessons A, E and B; but these activities had been of the shortest duration in Lesson C. In Lesson D, learners changed roles with the teacher, during repetition and verification. They spent a long time selecting cards to present to the class, and in formulating questions which were sometimes incorrect. Also, a very long time was employed in learning a song; all these could account for the long duration of activities in Lesson D.

The frequency of action and speech for each of the five lessons: A, 178, B, 254, C, 123, D, 183, E, 226. The new materials had been most frequently used by learners of Lessons B and E, and to a lesser degree by learners of Lesson D. But in Lessons A and C, they had been the least frequent.

And while action and speech had not been of long duration in Lesson B, they had been the most frequent; besides these were mostly speech responses. But, some classrooms took up a considerable length of time in action responses; for example, pupils' displacement in Lesson A, the selection of flash cards and the long time spent in formulating questions in Lesson D; also, learners of Lesson E spent quite some time drawing.

Language Use. The relative duration for each of the five lessons: A, 10.36 minutes, B, 7.24 minutes, C, 4.39 minutes, D, 10.94 minutes, E, 5.70 minutes. The duration of this set was comparatively longer in Lessons D and A than in Lessons B, E and C.

The frequency of language use for each of the five lessons: A, 372, B, 475, C, 303, D, 435, E, 373. Although the frequency of this behaviour was high for all five classrooms, it was predominantly used by the teachers in B and D.

Language had been used for a number of purposes which the writer has already presented in previous paragraphs. However, for the purpose of comparison, the most dominant behaviours will be mentioned. *Questioning* had been frequently used by all teachers except teacher D; in Lesson E, there were 136 questions asked by teacher, and 113 in Lesson A, 105 in Lesson B, 95 in Lesson C and the least of all, were 33 in Lesson D. But there were differences in the duration of this set; for example, questioning had been of long duration in Lessons A, C and E, whereas it was of shorter duration in Lessons B and D. *Iteration* was frequently used by teacher B, and was of long duration; but this behaviour was of shorter duration and frequency in Lessons A, C, D and E. Demonstration had been the most frequently used by teacher D and to a lesser degree by teachers A, B, C and E. The use of the mother tongue had been most frequently used by teacher B and was of long duration; it had been used by teachers C, D and E, infrequently and for a short duration. Behaviours of shorter duration and frequency were completing learners' statements, correction of pronunciation and explanations; these were observed in all five classrooms. Tables 23 and 24, as well as Figures 67 and 68 (b) illustrate the frequency and duration of language use.

Language Performance (individual). The relative duration of language performance for each of the five lessons: A, 2,08 minutes, B, 1.15 minutes, C, 1.88 minutes, D, 2.93 minutes, E, 4.38 minutes.

The relative duration of individual utterances (correct and incorrect) had been the longest in Lesson E, and the briefest in Lesson B. Whereas the duration of learners' utterances in Lesson D, A and C were in between these two extremes.

The frequency of language performance (individual) for each of the five lessons: A, 113, B, 78, C, 118, D, 103, E, 207. The frequencies shown are the totals of individual utterances, and these include correct and incorrect utterances. Learners in Lesson E responded to the highest number of utterances, but the least number of responses had been made in Lesson B. And in Lessons A and C, learners' utterances were equally frequent, but less so in Lesson D.

Therefore, the new didactic points occurred with the highest frequency and duration in Lesson E; whereas, they had been used most infrequently and for the shortest duration during Lesson B. However, learners in Lesson B made the least number of errors: (Lesson B, 28, Lesson A, 45, Lesson C, 45, Lesson D, 38 and Lesson E, 41); but pronunciation errors were the highest in Lesson E, there were 71 occurrences, and very low in the four other lessons. The highest number of pronunciation and grammar errors were made in Lesson A, and the highest number of mistakes, in all three types of errors, occurred in the utterances of learners in Lesson D. And these mistakes had been comparatively infrequent in the other lessons. The frequency and the duration for error types are shown on Table 19, and on Figures 29 (b), 30 (b).

The behaviour indicators of the five lessons A, B, C, D and E will now be compared.

1.2 *Indicators of Teaching Behaviours*

Indicators of quality can be used to distinguish between lessons. In order to do this, quantitative behaviour indicators of activity, variety, intensity and performance will be the measure used in the comparison. And since the behaviour indicator is an average or ratio, the indicator of highest numerical importance would demonstrate the degree of activity, variety, intensity and performance of one lesson as compared with the others. The comparison will be made in the following manner: First, the activity and variety of the total behaviours and the specific behaviours will be presented. Then the behaviour indicators of intensity and performance will follow.

Activity indicators for total behaviours in each of the five lessons: A, 19, B, 26, C, 18, D, 15, E, 18. The indicators show that Lesson B had been the most active, since the new didactic points were used 26 times every minute during the lesson. And, whereas the average use of the materials was lower in Lessons A, C, D and E, there were similarities of activity found in these four lessons.

Variety indicators for the total behaviours in each of the five lessons: A, 0.2763, B, 0.2770, C, 0.3500, D, 0.2425, E, 0.3112. In general all the lessons appear to have been equally varied with the exception of Lesson D. However, there was a degree of difference which was marginal among them. For instance, Lessons C and E appeared more varied than Lessons B and A, while Lesson D had been the least varied.

Therefore, Lesson B had been the most active but less varied, whereas, Lesson C had been the most varied; and lessons A, D and E had been less active as well as less varied.

TABLE 18

Sum of Duration and Frequencies of Behaviours Five Lessons

BEHAVIOURS	LESSON A		LESSON B		LESSON C		LESSON D		LESSON E	
	S. of DR	S. of FR.	S. of DR.	S. of FR.	S. of DR.	S. of FR.	S. of DR.	S. of FR.	S. of DR.	S. of FR.
Articulation	4.85	148	5.60	234	2.89	97	5.02	154	4.44	200
Presentation	0.41	17	0.00	000	1.10	11	5.10	184	1.77	56
Repetition	3.56	146	3.50	174	2.79	174	2.13	85	1.62	118
Verification	4.66	150	5.28	260	2.42	120	3.79	156	4.20	226
Didactic Actions	0.44	7	0.00	000	0.80	16	8.69	95	2.94	30
Speech thru Speech	0.89	47	0.24	14	0.23	17	0.04	2	0.56	44
Action & Speech	7.90	178	4.93	254	2.63	123	10.13	183	7.01	226
Language Use	10.36	372	7.24	475	4.30	302	10.94	435	5.70	373
Language Performance	2.08	113	1.15	78	1.88	118	2.93	103	4.38	207

S. of DR means Sum of duration. S of FR. means Sum of frequency.

TABLE 19

Frequencies of Language Use, in Five Lessons

LESSON	(1) LANGUAGE USE (frequency)	(2) CATEGORIES (No. of cat. in Lang. Use	(3) LANGUAGE 1 (sum of frequencies)	(4) TIME (in minutes)	A (i) MEANS (for LU 1/4)	A (ii) MEANS (cate. over time 2/4)	B (i) MEANS L1 over time 3/4)	B (ii) RATIO L1 over Lang Use 3/1)
A	372	7	000	10.36	35.91	0.6757	0000	0.0000
B	475	9	108	7.24	65.61	1.1050	15.0000	0.2274
C	302	9	27	4.30	70.23	2.0930	6.2791	0.0894
D	435	9	03	10.94	39.74	0.8223	0.2741	0.0069
E	373	8	02	5.75	64.81	1.3901	0.3475	0.0054

Cate. Categories; LU Language Use ; L1 Mother tongue.
1/4 means column one divided by column four.

FIGURE 29

Duration of Didactic and Language Behaviours in Five Classrooms

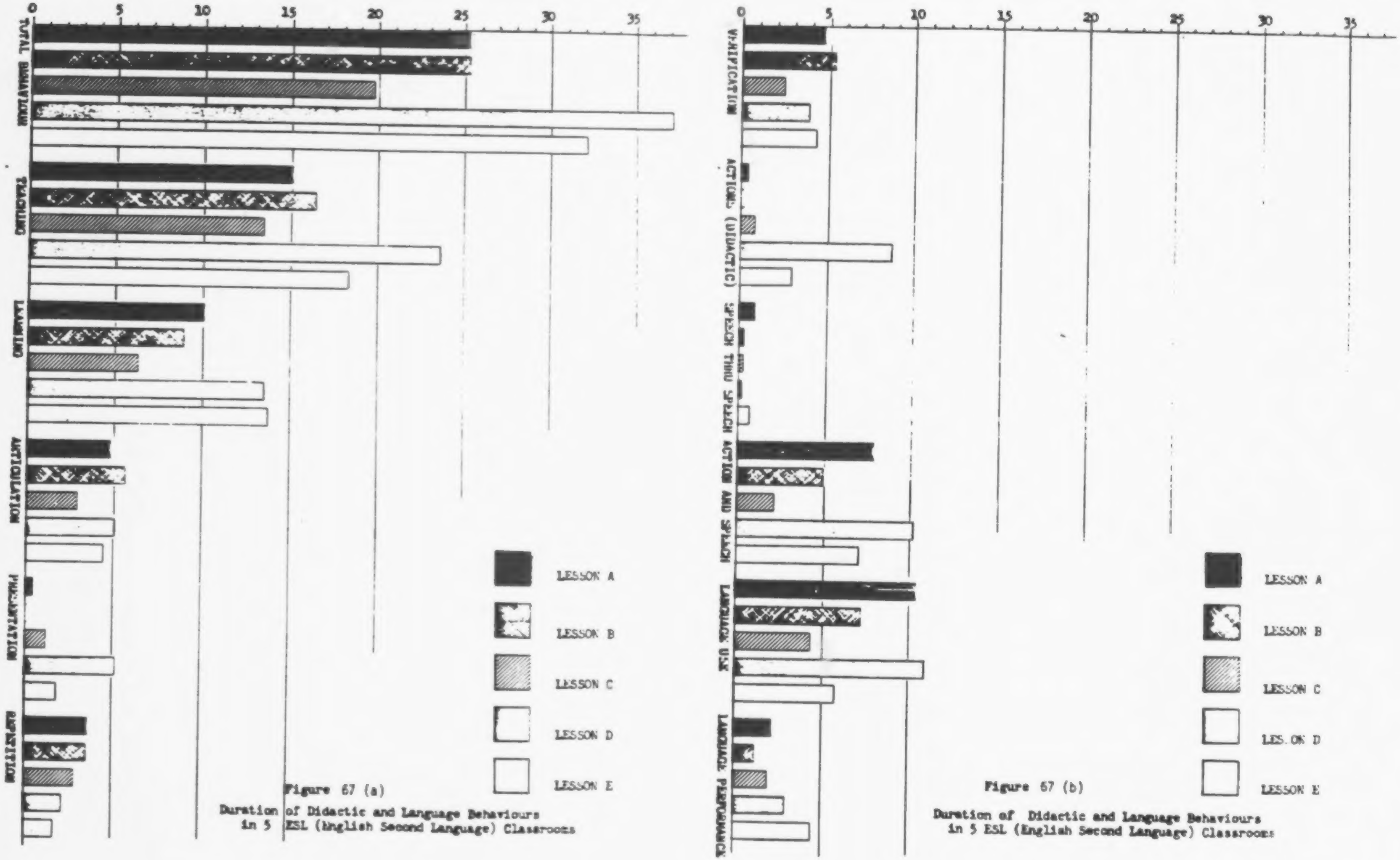


Figure 67 (a)

Duration of Didactic and Language Behaviours in 5 ESL (English Second Language) Classrooms

Figure 67 (b)

Duration of Didactic and Language Behaviours in 5 ESL (English Second Language) Classrooms

FIGURE 30

Frequency of Didactic and Language Behaviours in Five Classrooms

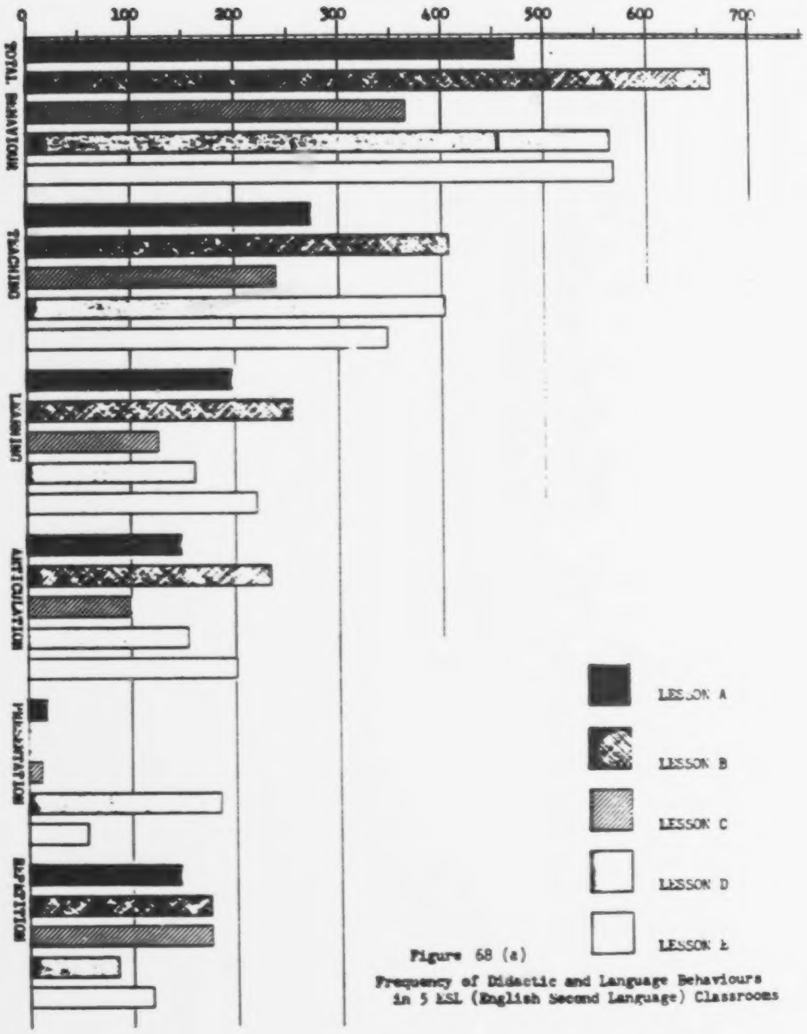


Figure 68 (a)

Frequency of Didactic and Language Behaviours in 5 ESL (English Second Language) Classrooms

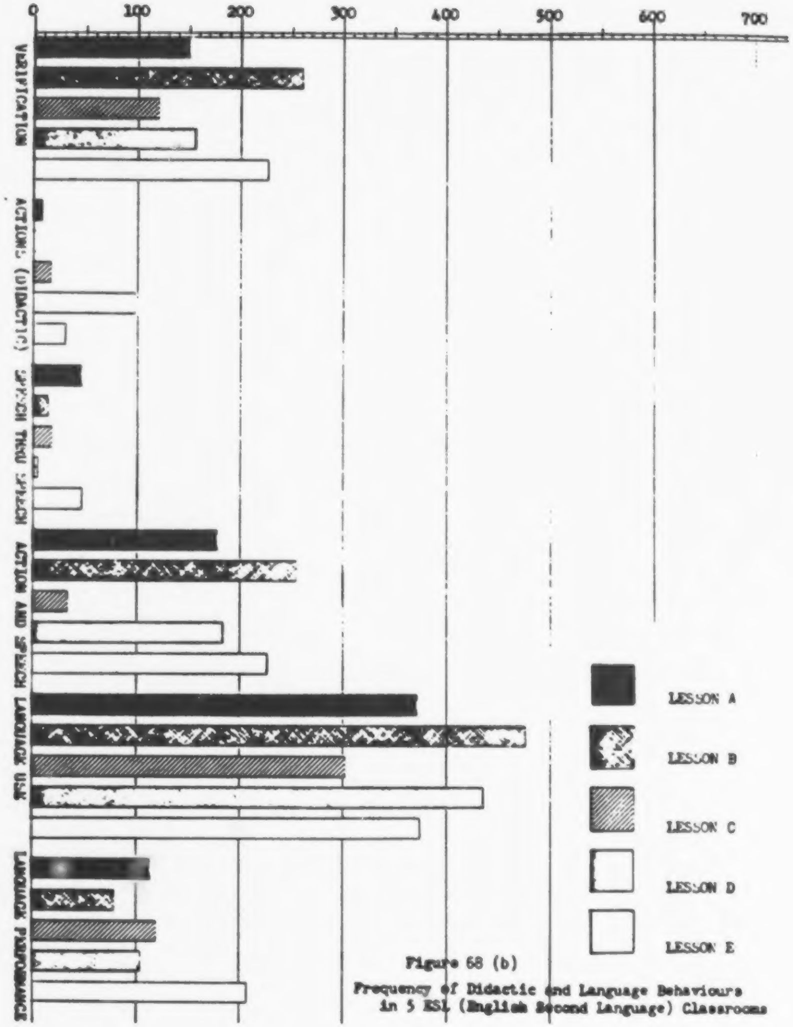


Figure 68 (b)

Frequency of Didactic and Language Behaviours in 5 ESL (English Second Language) Classrooms

Activity indicators for articulation in each of the five lessons: A, 31, B, 42, C, 34, D, 31, E, 45.04. Lessons B and E were similar, in that articulation was highly active in both; and whereas, this behaviour was used with equal frequency by teachers A and D, articulation had been the least active. Lesson C was in between these two extremes, and might be termed moderately active.

Variety indicators for articulation in each of the five lessons: A, 1.0417, B, 0.8929, C, 1.0381, D, 1.1952, E, 1.1261. The variety indicators show that there was one change made every minute in interest-promoting categories during articulation, and for all five classrooms. But there were marginal differences, Lessons D and E showed more variety than Lessons A, C and B. Therefore, the analyses of the five lessons show that articulation had been the most active and varied in Lesson D.

Activity indicators for presentation in each of the four lessons: A, 41.0, C, 10.0, D, 36.10, E, 32.0. Lesson A had been the most active of the four lessons; the new didactic points had been exposed to learners 41 times every minute during presentation in Lesson A; but less frequent in lessons D and E. And presentation was the least active in Lesson C.

Variety indicators for presentation in each of the four lessons: A, 7.31, C, 5.40, D, 1.60, E, 2.26. Lesson A had been the most varied; teacher A used about seven categories every minute during presentation. And presentation in Lesson C had been varied but the least variety during this set was in Lessons E and D. Therefore, presentation had been the most active and varied in Lesson A; and whereas, it had been very active in Lesson D, presentation was the least varied of the four lessons. And comparing Lessons E and C, there was more activity and less variety in E, but greater variety and less activity in C.

Activity indicators for repetition in each of the five lessons: A, 41.01, B, 50.00, C, 62.04, D, 40.00, E, 73.00. The new didactic points had been reinforced 73 times every minute in Lesson E, but were less frequently used minutely, in Lessons A, B, C and D. Teachers D and A made the least frequent use of repetition.

Variety indicators for repetition in each of the five lessons: A, 1.12340, B, 1.1429, C, 1.7921, D, 2.8200, E, 1.8520. Teacher D used about three techniques of repetition every minute, whereas, teachers E and C changed these techniques twice a minute, and approximately one change occurred every minute with teachers A and B.

Although repetition had been the most active in Lesson E, it lacked variety. But in Lesson D, repetition was the most varied, while least active of the five lessons.

Activity indicators for verification in each of the five lessons: A, 32.20, B, 49.24, C, 50.00, D, 41.16, E, 54.00. Teacher E checked learners more frequently every minute than teachers in the four other lessons. And, although teachers C and B checked learners' progress less frequently, this was done more infrequently by teachers D and A.

Variety indicators for verification in each of the five lessons: A, 1.5021, B, 1.3258, C, 2.4800, D, 1.6000, E, 1.4300. Teacher C changed interest-promoting categories more than twice every minute during verification. This change of categories was also high in lesson D, but teachers in Lessons A, B and E made about one category change every minute on an average, during verification.

While verification had been very active in Lesson E, it lacked variety. However, verification in Lesson C was the most varied of the five lessons, and activity, during checking of the new materials, was second to teacher.

Activity indicators for didactic actions in each of the five lessons: A, 16.00, C, 20.05, D, 11.00, E, 10.20. Teacher C used didactic actions, during the presentation of the new materials, approximately 20 times every minute. Therefore, learners in Lesson C had been taught meaning, twice times more often than learners in Lessons D and E. And teacher A's didactic actions were less frequent than teacher C's.

Variety indicators for didactic actions in each of the four lessons: A, 9.0910, C, 2.5063, D, 0.9206, E, 1.400. Teacher A used the highest number of action categories a minute; this was very varied compared with Lessons E and D, where teachers changed about one category every minute during didactic actions. However, didactic actions were moderately varied in Lesson C.

Therefore, teacher A's rate of didactic actions was active and varied; whereas didactic actions had been as active as they were varied in Lesson C. But the least activity and variety during didactic actions were observed in Lessons D and E.

Activity indicators for speech thru speech in each of the five lessons: A, 53.0, B, 58.33, C, 68.00, D, 50.00, E, 79.00. These indicators represent the rate at which speech drills were used every minute during the lesson. Therefore, teacher E used drills 79 times a minute, and this was the most active of the five lessons.

Variety indicators for speech thru speech in each of the five lessons: A, 1.1236, B, 4.1667, C, 8.000, D, 25.000, E, 0.6000. The categories changed very frequently in Lesson E, but this was the rate at which they were changed. And whereas, the categories changed less often in Lessons C, B and A, the least number of changes were made in Lesson E.

And, although speech thru speech had been very active in Lesson C, it was the least varied of the five lessons. And in Lesson D, these drills had been the most varied but least active.

Activity indicators for learners' action and speech activities in each of the five lessons: A, 22.53, B, 51.00, C, 16.00, D, 18.00, E, 32.24. Learners' didactic activities had been the liveliest in Lesson B, and somewhat less active in Lesson E. But activities had been the least active in Lessons C, D and A. Moreover, the new materials were used three times more often in Lesson B than in Lesson C, twice times more than in Lesson A and nearly three times more than learners in Lesson D.

Variety indicators for learners' action and speech activities in each of the five lessons: A, 0.8861, B, 1.4000, C, 1.9600, D, 0.5923, E, 1.1412. Learners' didactic activities had been most varied in Lesson C, and to a lesser degree in Lesson B. However, Lessons E and A had been equally varied, while activities in Lesson D showed the least variety.

Therefore, learners' activities had been highly active, but moderately varied in Lesson B; whereas, in Lesson C didactic action and speech had been the least active but most varied of the five lessons. And, although learners' activities were active in Lesson E, they lacked variety. But, they were neither active nor varied in Lessons A and D.

The behaviour indicators of intensity and performance will be compared in the following paragraphs, since the indicators of activity and variety have been compared above.

Intensity indicator in each of the five lessons: A, 8.00, B, 10.00, C, 6.35, D, 4.30, E, 7.10. The data used to calculate these behaviour indicators are shown on Table 25. The interaction of teacher/learner showed the highest intensity in Lesson B; there were 10 such interactions with teacher every minute, throughout the entire lesson. Learners in Lesson A used the new materials less densely during didactic activities, the same was true in Lesson E and C. But in Lesson D, learners interacted the least with the teacher, for the entire duration of the lesson.

Performance indicators will be compared in the following manner:

A (i) Indicators for correct utter, (ind.) per minute of alloted time
A (ii) Indicators for ratio of correct utterances
A (iii) Indicators for errors (pronunciation, vocabulary and grammar)
A (iv) Indicators for errors (vocabulary and grammar)

TABLE 20

Intensity Indicators for Five Lessons

LESSON	ACTIONS (individual	ACTIONS (group)	ACTION AND SPEECH(ind.)	UTTERANCES (ind.)	UTTERANCES (group)	TOTAL (Action & Speech)	TIME (in mins)	INTENSITY INDICATOR
A	5	26	00	113	53	197	25.3	7.80
B	1	00	00	78	166	245	25.3	9.68
C	2	00	00	118	7	127	20,0	6.35
D	38	00	39	103 64	17	158	37.0	4.30
E	3	14	18	207 189	3	227	32.1	7.10

A (i). Indicators for correct utterances (ind.) per minute of the alloted time, in each of the five lessons: A, 17, B, 19.13, C, 19.00, D, 6.00, E, 17.25. The rate of correct individual utterances every minute had been the highest in Lesson B and C, and less frequent in Lessons A and E. But learners' rate of success was very low in Lesson D.

A (ii). Indicators for ratio of correct utterances in each of the five lessons: A, 0.310, B, 0.280, C, 0.320, D, 0.175, E, 0.333. Learners in all the classrooms with the exception of Lesson D, made the same number of correct responses. Nevertheless, there were marginal differences among them. For instance, 33 percent of all individual utterances were correct in Lesson E, and in Lesson C, 32 percent of all individual utterances were correct; 31 percent were correct for Lesson A and 28 percent were correct in Lesson B. And least of all, 17 percent of all individual utterances were correct in Lesson D.

A (iii). Indicators for error types in each of the five lessons: A, 0.60, B, 0.60, C, 0.54, D, 0.70, E, 0.54. The proportion of errors was highest in Lesson D and to a lesser degree in Lessons A and B. The lowest ratio was in Lessons C and E.

A (iv). Indicators for errors of vocabulary and grammar, in each of the five lessons: A, 0.40, B, 0.36, C, 0.36, D, 0.37, E, 0.20. The ratio of errors of vocabulary and grammar seemed equal in Lessons A, B, C and D but not in Lesson E. The difference among learners may be shown as percentages. For instance, 40 percent of mistakes in Lesson A were in vocabulary and grammar; 36 percent were found for Lessons B and C equally, and 37 percent of mistakes were made in Lesson D. But only 20 percent of errors were made in Lesson E, in vocabulary and grammar.

In sum, the comparison of the five lessons shows that Lesson B had been the most active and intense, whereas, Lesson C had been the most varied. But Lesson D had been the least active and least varied of all five lessons. And while Lessons A and E had been less active than B, Lessons A, E and B were equally varied. Besides, there were seven didactic behaviours identified for Lessons A, C, D and E, but only five could be identified for Lesson B. In Lesson E six behaviours had been active and five were varied. On the contrary, all five behaviours in Lesson B had been active though less varied. And whereas, five behaviours in Lesson C had been active, six were comparatively varied. However, there were only three behaviours found active in Lesson A and four were varied. And two behaviours were active in Lesson D, while five were varied. The performance of individual learners showed that the highest successes had been in Lessons B and C, and to a lesser degree in Lessons E and A, but the least success appeared in Lesson D. Similarly, the highest ratio of errors was found in Lesson D, but it was lower in lessons A and B, and even lower still in Lessons C and E.

2. Relating Teachers to Learning

We are now in a position to study the relationship between teaching behaviour and learning. To what extent was the linguistic performance of the 11-year-olds, in the beginning class of English as a second language, due to the use of didactic behaviours. The variables used to test this hypothesis consisted of eleven concrete and two structure words These new materials were taught in lesson samples, and they were also used in the construction of tests (listening/speaking). Hence, didactic behaviours in the form of lessons were recorded and analysed; and linguistic performance was judged on learners' success on the tests. And, in order to find out what effect, if any, the didactic behaviours had on learners' linguistic performance, Pearson's simple correlation was used to calculate coefficients of correlation.

In addition, the five classrooms were divided into two groups. This dichotomy was made for two reasons. First, Scheffé's comparison of the mean test scores showed that Lesson A was different from Lesson E on the listening comprehension test. Secondly, the lesson analysis showed that teacher A omitted to teach the new concrete words (they were used during the lesson), but presented only the two structure words to learners of Lesson A. Because of these reasons, Lessons A, B, C, D and E were labelled Group One, and Lessons B, C, D and E, Group Two.

The relationship will be presented and interpreted as correlation coefficients: 1. for the overall behaviour and tests, 2. for the specific didactic behaviours and tests and 3. for language use and tests.

2.1 *Overall Classroom Behaviours*

Activity Indicator/Listening	Group One:	r = 0.40;	Group Two:	r = 0.52
Activity Indicator/Speaking	Group One:	r = 0.70;	Group Two:	r = 0.74

The high positive correlations found between an active lesson and performance on the listening comprehension test, in Groups One and Two, were surpassed by the even higher correlations between an active lesson and the speaking test results. Therefore, it appears that learners tended to be very successful on the listening and speaking tests, as a result of the increased use of the didactic points every minute, during teaching and learning.

Variety Indicators/Listening	Group One:	r = -0.20;	Group Two:	r = -0.32
Variety Indicators/Speaking	Group One:	r = 0.61;	Group Two:	r = 0.60

The correlations between a lesson that was varied and listening comprehension were low and negative in Group One, but high and negative in Group Two. This seems to indicate that a lesson that was less varied might have brought about good results on the listening comprehension of Group One, but too much variety did not in Group Two. However, the high positive correlations, between a varied lesson and the speaking test suggest, that as interest-promoting categories increased every minute, during the lesson, learners might be expected to perform more successfully on the speaking test.

Intensity Indicators/Listening Group One: r = 0.33; Group Two: r = 0.30
Intensity Indicators/Speaking Group One: r = 0.53; Group Two: r = 0.72

The correlation coefficients which were found between lessons which were intense and the tests (listening/speaking) were very high and positive. It would seem that learners' close contact with the new didactic points during learning, might have brought about a high incidence of success on their linguistic performance (listening and speaking tests), in both Group One and Group Two.

Performance Indicators/Tests (listening/speaking)

A (i) Correct utterances (ind.) per minute/Listening
Group One: r = 0.15; Group Two: r = 0.10
A (i) Correct utterances (ind.) per minute/Speaking
Group One: r = 0.70; Group Two: r = 0.85

The correlations found between learner's correct utterances and the listening test were low but positive. This suggests that when the average number of correct utterances decreased during the lesson, learners might be expected to perform less well on listening comprehension. But the very high indices found between correct individual utterances and speaking suggest, that the scores on the speaking test would increase, as learners' correct utterances increased every minute during the lesson.

A (ii) Proportion of correct utterances (ind.)/Listening
Group One: r = -0.05; Group Two: r = 0.21
A (ii) Proportion of correct utterances (ind.)/Speaking
Group One: r = 0.42; Group Two: r = 0.70

The correlation found between the proportion of correct individual utterances and listening in Group One was low and negative. This also suggests a very weak relationship, almost zero, which might have unpredictable results. But the low and positive index for A (ii) and listening in Group Two suggests, that a small proportion of correct utterances during the lesson might have caused similar low results on the listening test. But the correlations between the proportion of correct utterances and speaking were high and positive in both groups. This might be interpreted to mean that a high proportion of correct utterances during the lesson, might have caused a correspondingly high degree of success on the speech performance of both groups.

A (iii) Incorrect utterances (ind.)/Listening
Group One: r = 0.11; Group Two: r = 0.30
A (iii) Incorrect utterances (ind.)/Speaking
Group One: r = -0.50; Group Two: r = -0.60

The correlation coefficients between the ratio of incorrect utterances and listening were positive in both groups. But the low index of Group One suggests that learners tended to have depressed scores on listening, as a result of the low ratio of correct utterances during learning. And the high index of Group Two suggests that listening scores increased as the ratio of correct utterances during the lesson increased. However, the correlations between the incorrect utterances and speaking were high but negative, and for both Groups One and Two. This could be interpreted to mean, that learners made fewer errors on the speaking test as a result of the high proportion made during learning. This applies to Group One as well as Group Two.

A (iv) Incorrect utterances (ind.) in vocabulary and grammar/Listening
Group One: r = 0.91; Group Two: r = 0.92
A (iv) Incorrect utterances (ind.) in vocabulary and grammar/Speaking
Group One: r = 0.02; Group Two: r = 0.30

The correlation coefficients between the ratio of errors in vocabulary and grammar and the listening comprehension test were high and positive for Group One and Group Two. This suggests that the high ratio of errors of vocabulary and grammar during learning, might have caused an increase of errors on the listening comprehension test. This could also apply to the correlation of A (iv) and speaking of Group Two, since the relationship between the two variables was close and positive. However, the index found for Group One demonstrates no relationship between the variables, therefore, results are unpredictable.

Tables 21 and 22 show the correlations found for the general behaviour indicators and the tests for the two groups, Group One and Group Two.

2.2 *Didactic Behaviours*

For the sake of simplicity, the three major categories (articulation, presentation, repetition), and sub-categories (verification, didactic actions, speech thru speech and action and speech) of the lessons were grouped under the label *specific didactic behaviours.* And the behaviour indicators of activity and variety were derived from the second analyses of the five lessons. Correlations were made between these behaviour indicators (activity and variety) and the listening and speaking test data. The order of presentation of the correlation coefficients will be as follows: 1. The correlations found for activity of the specific behaviours, first for the listening test and then for the speaking test. 2. The correlations found between the variety indicators of the specific behaviours for the listening test, and the speaking test.

2.2.1 Activity Indicators/Tests

A.I. of Articulation/Listening Group One: r = -0.61; Group Two: r = -0.50
A.I. of Articulation/Speaking Group One: r = 0.40; Group Two: r = 0.23

The correlations between an active articulation and results on the listening test were high but negative. This means that learners might have done poorly on listening comprehension, as a result of too much activity during articulation. But, the same was not true for activity and speaking; in Group One frequent use of articulation a minute, appears to have brought about desired good results in recognition of the new materials; in Group Two depressed scores on listening might have been caused by less activity during articulation.

A.I. of Presentation/Listening Group One: r = 0.10; Group Two: r = -0.60
A.I. of Presentation/Speaking Group One: r = - .97; Group Two: r = -0.98

The degree of relationship, between an active presentation and listening comprehension, was different in Groups One and Two. And, whereas infrequent presentation of the new didactic points in Group One, might have caused depressed scores on the listening test; a very active presentation, on the other hand, appeared to have caused the same results in Group Two. However, the correlations between an active presentation and speaking had similar results in both Groups One and Two. And, the relationship, between frequent presentation of the new materials every minute, and learners' performance on the speaking test, had been close but negative. This might be interpreted to mean that a highly active presentation would not bring about the desired behaviour change in learners' linguistic performance, that is to say, listening comprehension and speaking skills.

A.I. of Repetition/Listening Group One: r = -0.72; Group Two: r = -0.62
A.I. of Repetition/Speaking Group One: r = 0.41; Group Two: r = 0.23

TABLE 21

Correlation Indices: Behaviour Indicators and Oral Tests

GENERAL BEHAVIOURS / TEST		GROUP ONE		GROUP TWO	
		LISTENING	SPEAKING	LISTENING	SPEAKING
1.	Activity Indicators	0.40	0.70	0.52	0.74
2.	Variety Indicators	-0.20	0.61	-0.32	0.60
3.	Intensity Indicators	0.33	0.53	0.27	0.72
4.	Performance Indicators				
A (i)	Correct individual utterances per minute of allotted time.	0.15	0.70	0.10	0.85
A (ii)	Proportion of correct individual utterances out of total utterances (ind.)	-0.05	0.42	0.21	0.66
A (iii)	Incorrect individual utterances (pronunciation, vocabulary, grammar)	0.11	-0.50	0.26	-0.60
A (iv)	Errors in vocabulary and grammar (individual)	0.91	0.02	0.92	0.30

TABLE 22

Correlation Indices: General Behaviour Indicators and the Oral Tests

TEST / GENERAL BEHAVIOURS	LISTENING		SPEAKING	
	G1 (A,B,C,D,E)	G2 (B,C,D,E)	G1 (A,B,C,D,E)	G2 (B,C,D,E)
1. Activity Indicators	0.40	0.52	0.70	0.74
2. Variety Indicators	-0.20	-0.32	0.61	0.60
3. Intensity Indicators	0.33	0.30	0.53	0.72
4. Performance Indicators				
A (i) Correct individual utterances per minute of allotted time	0.15	0.10	0.70	0.85
A (ii) Proportion of correct individual utterances out of total utterances (ind.)	-0.05	0.21	0.42	0.70
A (iii) Incorrect utterances (ind.) (pronunciation, vocabulary and grammar errors)	0.11	0.30	-0.50	-0.60
A (iv) Errors in vocabulary and grammar (individual)	0.91	0.92	0.02	0.30

The correlation found for an active repetition and the listening test were very high but negative in both groups. This means that over reinforcement of the new didactic points might have caused more depressed scores on the listening comprehension test in Groups One and Two. But, the relationship between activity during repetition and speaking, was close and positive in Group One, but low and positive in Group Two. The correlations show that frequent repetitions every minute might have caused learners' scores to increase on the speaking test in Group One, but, fewer repetitions might have caused depressed scores on the speaking test of Group Two. Therefore, it would appear that less repetition might be predictive of listening success, whereas, frequent repetition might bring about success in speaking.

A.I. of Verification/Listening	Group One: r = -0.70;	Group Two: r = -0.42
A.I. of Verification/Speaking	Group One: r = 0.62;	Group Two: r = 0.50

The correlations between an active verification and listening were high but negative in both groups. This could be interpreted to mean that frequent checking every minute, during the lesson, did not bring about the desired behaviour change in learners' performance on the listening comprehension test. However, the correlations between a varied verification and speaking were high and positive. This means that learner's ability to utilize the new didactic points, independent utterances, might have increased as a result of frequent checking every minute, during teaching and learning. Therefore, variety during verification appeared to have brought about the desired behaviour change in learners' speech performance but not in their listening comprehension performance.

A.I. of Didactic Actions/Listening	Group One: r = 0.70;	Group Two: r = 0.73
A.I. of Didactic Actions/Speaking	Group One: r = 0.72;	Group Two: r = 0.92

The correlation coefficients found between activity during didactic actions and the listening and the speaking tests were very high and positive. This indicates that the relationship between frequent didactic actions, during presentation, and learners' linguistic performance, listening and speaking, had been close. Therefore, it appears that frequent use of didactic actions every minute during presentation, might have brought about the desired behaviour change in learners' linguistic performance (exemplified on the listening and speaking tests).

A.I. of Speech thru Speech/Listening Group One: $r = -0.72$; Group Two: $r = -0.65$
A.I. of Speech thru Speech/Speaking Group One: $r = 0.40$; Group Two: $r = 0.20$

The correlations for activity during speech drills and listening were high but negative in both groups. This might mean that very frequent drills, every minute during the lesson might have caused learners to perform less successfully on the listening comprehension test. However, the correlation between active drills and speaking was high and positive in Group One, but low and positive in Group Two. And whereas frequent drills every minute might have brought about increased scores in Group One, fewer drills might have caused more depressed scores on the speaking test in Group Two. Therefore, it appears that a very active speech thru speech might have brought about the desired linguistic performance on the speaking test but not on the listening comprehension test.

A.I. of Action and Speech/Listening Group One: $r = 0.069$; Group Two: $r = 0.48$
A.I. of Action and Speech/Speaking Group One: $r = 0.989$; Group Two: $r = 0.987$

The correlations found between activity during action and speech and the listening test were high and positive in Group Two but low and positive in Group One. This could be interpreted to mean that frequent use of the new didactic points during learning might have caused learners' scores to increase on the listening comprehension test, in Group Two, but infrequent use of the new materials during activities might have caused fewer successes in Group One. However, the correlations between an active action and speech and the speaking test were very high and positive. The relationship in the latter case was almost perfect. It appears that frequent didactic activities during learning were closely related to success on the speaking test. Therefore, learners' highly active participation during the lesson might have brought about the desired behaviour change in their linguistic performance which was reflected on the listening and speaking tests.

Table 23 show the correlation coefficients which were calculated between the activity indicators of the specific behaviours and the listening comprehension and the speaking tests.

2.2.2 Variety Indicators/Tests

V.I. of Articulation/Listening Group One: $r = -0.54$; Group Two: $r = -0.61$
V.I. of Articulation/Speaking Group One: $r = -0.75$; Group Two: $r = -0.90$

The correlations between articulation which was varied and the listening comprehension test were high and negative. But the correlations between this behaviour indicator and speaking far surpassed that of listening, and they were also negative. This could be interpreted to mean, that articulation which was too varied might not have brought about the desired change in learners' linguistic performance.

V.I. of Presentation/Listening Group One: $r = 0.81$; Group Two: $r = 0.55$
V.I. of Presentation/Speaking Group One: $r = 0.23$; Group Two: $r = 0.99$

The correlation coefficients found for a varied presentation and the listening test were high and positive in both groups, but the index of Group One showed a closer bond between variety during presentation and the listening test. And the correlations between a varied presentation and the speaking test were high and positive in Group Two, but low and positive in Group One. Moreover, the very high correlation of the behaviour indicator and speaking in Group Two seems to indicate that presentation that was varied might have brought about the desired change in learners' linguistic performance.

TABLE 23

Correlation Indices: Activity Indicators and Oral Tests

SPECIFIC BEHAVIOURS (ACTIVITY) / TEST	GROUP ONE		GROUP TWO	
	LISTENING	SPEAKING	LISTENING	SPEAKING
1. Articulation Indicators	-0.61	0.40	-0.46	0.23
2. Presentation Indicators	0.10	-0.97	-0.57	-0.98
3. Repetition Indicators	-0.72	0.41	-0.42	0.23
4. Verification Indicators	-0.70	0.62	-0.42	0.47
5. Didactic Actions Indicators	0.70	0.72	0.73	0.92
6. Speech thru Speech Indicators	-0.72	0.40	-0.65	0.20
7. Action and Speech Indicators	0.07	0.989	0.48	0.987

V.I. of Repetition/Listening Group One: r = -0.50; Group Two: r = -0.30
V.I. of Repetition/Speaking Group One: r = -0.42; Group Two: r = -0.90

The coefficients for a varied repetition and the listening test were high but negative in both groups. This could mean that too many changes in repetition techniques, every minute, might have caused learners to score less on the listening test. But the correlations between a varied repetition and the speaking test were also high and negative in both groups, but the relationship was far surpassed in Group Two.

V.I. of Verification/Listening Group One: r = 0.11; Group Two: r = 0.30
V.I. of Verification/Speaking Group One: r = -0.41; Group Two: r = 0.40

The correlations between verification and listening were high and positive in Group Two, but low and positive in group One. This might be interpreted to mean, that frequent changes every minute, of interest-promoting categories during verification, might have caused learners' scores on listening to increase in Group Two; whereas less variety in checking might have caused fewer successes in Group One. However, the correlations for a varied verification and speaking were strong and negative in Group One, but strong and positive in Group Two. This suggests that while variety during verification might have caused learners' success in Group Two, too much variety had adverse effects on learners' speaking scores, in Group One. Therefore, variety during verification might have caused the desired linguistic performance of learners in Group Two (listening and speaking test), but not in the speaking performance of learners in Group One.

V.I. of Didactic Actions/Listening Group One: r = 0.80; Group Two: r = 0.50
V.I. of Didactic Actions/Speaking Group One: r = -0.25; Group Two: r = 0.998

The correlations between variety during didactic actions and listening were high and positive in both groups. This suggests that frequent changes in interest-promoting categories, might have caused increased scores on the listening test. Whereas the correlation was very high and positive for varied actions and speaking in Group Two, it was low and negative in Group One. Moreover, this could mean that a high percentage of success in Group Two might have been due to variety during didactic actions, but the same results for Group One might have been caused by less variety in actions. Therefore, much variety during didactic actions seemed to have brought about the desired change in learners' linguistic performance, with the exception of speaking results in Group One.

V.I. of Speech thru Speech/Listening Group One: r = 0.10; Group Two: r = 0.20
V.I. of Speech thru Speech/Speaking Group One: r = -0.33; Group Two: r = -0.63

The correlations between a varied speech thru speech and listening were low and positive for both groups. It appears that drills which were not varied might have caused scores to be depressed. However, the correlations for variety during drills and speaking were high but negative. This might be interpreted to mean that too much variety during drills might have caused depressed scores on the speaking test as well. Therefore, it would appear that a less varied speech thru speech did not bring about the desired change in learners' linguistic performance (listening test performance) and the frequent category changes during drills also had the same effect on the speaking test.

V.I. of Action and Speech/Listening Group One: r = 0.03; Group Two: r = 0.28
V.I. of Action and Speech/Speaking Group One: r = 0.94; Group Two: r = 0.945

The correlation coefficients found for variety during action and speech and the listening test were low and positive in Group One, but high and positive in Group Two. This could be interpreted to mean that learners' active participation during learning and teaching, might have caused good results on their listening comprehension test scores; but less variety might have caused less success in group One. However, the correlations for varied participation and the speaking test were extremely high and positive. This could mean that frequent use of the new didactic points during learning activities might have caused their scores on the speaking test to increase. Therefore, variety during learners' action and speech activities appeared to have brought about the desired change in their linguistic performance, particularly on the speaking test which showed almost perfect correlation. Table 24 shows the correlations for the variety behaviours (specific) and the tests results.

2.3 *Language Use*

The writer will attempt to present and interpret the correlation coefficients which were found for the teachers' language behaviour and the learners' tests performance. There were two aspects of teachers' language behaviour which were studied: A *language use* and B *language one* (mother tongue). The correlation will be made in the following manner:

A (i) The occurrence of language use per minute of alloted time and the tests (listening and speaking). A (ii) Categories used per minute during the time alloted to language use and the tests. B (i) Occurrence of L_1 per minute of alloted time of language use, and B (ii) Proportion of L_1 to the total frequency use and the tests. These correlations were calculated for Group One and Group Two.

A (i) Occurrence of Lang. Use/List. Group One: r = -0.40; Group Two: r = 0.05
A (i) Occurrence of Lang. Use/Speak. Group One: r = 0.90; Group Two: r = 0.83

The correlation between teachers' frequent use of language, every minute, and the listening test was high and negative in Group One, but low and positive in group Two. This might be interpreted to mean that too many occurrences of language, every minute, might have caused fewer scores for Group One learners. And the relationship between the variables in Group Two was unpredictable. However, the correlations between language used frequently, and the speaking test were high and positive. It would appear that teachers' frequent use of language was closely related to the high scores obtained on the speaking test in Group One and Group Two alike. Therefore, frequent use of language by teacher might have brought about the desired change in learners' linguistic performance on the speaking test but not on the listening test.

A (ii) Categories used perminute of alloted time/Listening
Group One: r = -0.30; Group Two: r = 0.06
A (ii) Categories used per minute of alloted time/Speaking
group One: r = 0.70; Group Two: r = 0.61

The correlation between A (ii) and listening was strong but negative in Group One but low and positive in group Two. This indicates that learners tended to obtain lower scores on the listening test as a result of too many technique changes during teachers' language use; but in Group Two, learners scored less, as a result of fewer category changes in language use. Therefore, while less category changes were needed in Group One, more were required in Group Two, in order to obtain desirable results in both groups.

Nevertheless, the correlations were high and positive between A (ii) and the speaking test. This might be interpreted to mean that learners' scores increased on the speaking test as a result of frequently used techniques of language use. Therefore, it appears that frequent changes in techniques of language use might have brought about the desired change in learners' linguistic performance on the speaking test, but not on the listening test of Group One. While in Group Two, fewer techniques caused the same results.

Since the mother tongue had been used only in Group Two, no comparison was made with Group One.

B (i) Occurrence of Language One/Listening Group Two: r = 0.74
B (i) Occurrence of Language One/Speaking Group Two: r = 0.82

TABLE 24

Correlation Indices: Activity Indicators and Oral Tests

SPECIFIC BEHAVIOURS (VARIETY) / TEST	GROUP ONE LISTENING	GROUP ONE SPEAKING	GROUP TWO LISTENING	GROUP TWO SPEAKING
1. Articulation Indicators	-0.54	-0.75	-0.61	-0.896
2. Presentation Indicators	0.81	0.23	0.55	0.99
3. Repetition Indicators	-0.50	-0.42	-0.28	-0.86
4. Verification Indicators	0.11	-0.41	0.29	0.37
5. Didactic Actions Indicators	0.80	-0.25	0.46	0.998
6. Speech thru Speech Indicators	0.10	-0.33	0.20	-0.63
7. Action and Speech Indicators	0.03	0.94	0.28	0.94

The correlations between the frequent use of the mother tongue and the listening and the speaking tests were high and positive. This might be interpreted to mean that learners might be expected to score high on both tests as a result of frequent use of the mother tongue every minute during teachers' use of language. Therefore, the frequent use of the mother tongue appeared to have brought about the desired change in learners' linguistic performance.

B (ii) Proportion of L_1 to Language One/ListeningGroup Two: $r = -0.74$
B (ii) Proportion of L_1 to Language One/Speaking Group Two: $r = 0.82$

The correlations between B (ii) and the listening comprehension and the speaking tests were very high and positive. This might be interpreted to mean that learners' close contact with the mother tongue during teaching/learning, brought about a high incidence of success on the tests. Therefore, it appears that when a high percentage of L_1 was used during the lesson, this might have brought about the desired change in learners' linguistic performance.

Table 25 shows the correlations for Language Use: A (i) and A (ii) and Language One B (i) and B (ii).

2.4 *Summary of Correlations*

Correlation coefficients had been calculated between the behaviour indicators (activity, variety, intensity and performance) and the mean gain scores obtained on the listening comprehension and the speaking tests, in order to find out what effects, if any, these behaviours had on the linguistic performance of learners, in a beginning class of English as a second language. The results show that very active and intense lessons might have brought about the desired changes in learners' linguistic performance. And whereas a varied lesson had the same effects on learners' speech performance in both groups, less variety had equally good results on the listening performance of learners in Group One, and too much variety did not bring about the desired effects in Group Two. The high correlations show that there was a close relationship between learners' successful utterances during the lesson and their performance on the speaking test; and low scores on the listening test might have been caused by less frequent success during learning. This interpretation might also apply to high positive correlations found for the proportion of correct utterances and speaking, and the low positive correlation for listening in Group Two. But, in Group One the effects were unpredictable. Nevertheless, the high proportion of errors made during learning might be linked to fewer errors on the speaking test, but might have caused learners' errors to increase on the listening test in Group Two; whereas less errors were made on the listening in test Group One as a result of fewer errors during learning. And the high proportion of errors of vocabulary and grammar which were made in both groups, during learning, might have caused similar errors in learners' performance on the listening and speaking tests.

The correlation coefficients show that too much activity and variety during articulation did not bring about the expected change in learners' performance on the listening comprehension test, and the same was true for too much variety and its effects on learners' speech performance. But, an active articulation appeared to have brought about the expected change on the speaking test. Presentation that was active and varied seemed to have good effects on the listening comprehension test, whereas, too much activity did not bring about good results on the speaking test, while a varied presentation did. Whereas too much activity and variety appeared ineffective on the listening test, too much variety brought about the same results on the speaking test, but, a very active repetition appeared to have been effective. And although activity during verification appeared to have been effective on the speaking but not on the listening test, variety during verification was effective in Group Two, both on the listening and the speaking tests. Besides, in Group One less variety caused depressed scores on the listening test, while too much variety during verification might have had adverse effects on speaking. Didactic actions which were highly active and varied appeared to have brought about the desires change in learners' linguistic performance in all groups, with the exception

TABLE 25

Correlation Indices: Language Use and Oral Tests

TEST / LANGUAGE USE	GROUP ONE LISTENING	GROUP ONE SPEAKING	GROUP TWO LISTENING	GROUP TWO SPEAKING
A (i) Occurrence of language use (per minute of allotted time)	-0.40	0.90	0.05	0.83
A (ii) Types of language changed per minute of allotted time	-0.30	0.70	0.06	0.61
B (i) Occurrence of L1 per minute of allotted time	0.74	0.84	0.74	0.54
B (ii) Proportion of L1 to total frequencies of language use	0.74	0.62	0.74	0.82

of Group One, it appears that less variety might have brought about equally good results on the speaking performance of learners in that group. Speech thru speech which was too active but not varied seemed to have been ineffective on the listening test. Whereas great activity during speech thru speech had been very effective on the speaking test, too much variety was not. Learners' action and speech activities which were very active and varied appear to have brought about the desired changes in their linguistic performance on the speaking test in both groups, and for Group Two on the listening comprehension test. Less activity and variety during learners' activities in Group One, might have caused them to score poorly on the listening test.

The correlations found for the teachers' language behaviour reveal that frequent language use, and too many category changes, every minute, had not been effective in bringing about the desired change in learners' linguistic performance on the listening test, in Group One, and in Group Two correlations were unpredictable. However, it seems that frequent use of language and the frequent category changes every minute might have been effective on the speaking test, and in both groups. Furthermore, the high positive correlations found for the frequent use of the mother tongue during teaching, and the performance on the listening and the speaking tests, seem to indicate that learners' close contact with the mother tongue during learning might have caused a high incidence of success on both the listening comprehension and the speaking tests (See Tables 26 and 27).

3. Interpretation of Teaching and Learning Differences

The comparison of the analyses of the five lessons revealed that Lesson D had been of the longest duration but behaviours were of lesser duration than in Lesson B which had the highest frequencies. Lessons A and B were of the same duration and Lesson E had been almost as long as Lesson D; during that time the new materials were repeated with equal frequency in both. Lesson C had been of the shortest duration and frequency. The comparison of the results of the qualitative analyses of the five lessons showed that Lesson B had been the most active and intense, but less varied than Lesson C. And Lesson D had been the least active, varied and intense of the five lessons. But, in Lessons A and E behaviours had been more active, varied and intense than in Lesson D. And, while Lesson C had been the most varied, it was neither active nor varied. The performance indicators showed that learners in Lessons B and C had obtained the highest averages for correct utterances, but this was lower in Lessons E and A, and the lowest average was found in Lesson D. Likewise, the percentage of errors had been the highest in Lesson D, and to a lesser degree in Lessons A and B, but the percentage of errors made by learners in Lessons C and E had been the lowest.

Activity and variety indicators were also calculated for the specific behaviours (articulation, presentation, repetition, verification, action and speech, speech thru speech and action and speech). These behaviour indicators will be placed in descending order of importance, in the following presentation.

Articulation had been very active in Lessons E and B and to a lesser degree in Lessons C, A, D; but Lesson D had been the most varied, while Lesson B showed the least variety. And Lessons E, A and C were in between these two extremes.

Presentation showed great activity in Lesson A, D, E and C; and the widest variety was in Lessons A, D, E and C.

Repetition had been very active in Lessons E, C, B, A and D. But while it had been the least active, repetition was the most varied in Lesson D and least varied in Lessons A, B, C and E.

Verification had been very active in Lessons E, C, B, D and A; and the most variety had been in Lessons C, D, A, E and B.

TABLE 26

Summary Table of Correlations

TEST / BEHAVIOURS	LISTENING		SPEAKING	
	GROUP ONE	GROUP TWO	GROUP ONE	GROUP TWO
Activity	0.40	0.52	0.70	0.74
Variety	-0.20	-0.32	0.61	0.60
Intensity	0.33	0.30	0.53	0.72
Performance:				
Correct per minute	0.15	0.10	0.70	0.85
Ratio corr. utter.	-0.05	0.21	0.42	0.70
Ratio incorrect utt.	0.11	0.30	-0.50	-0.60
Ratio voc. & gramm. (incorrect)	0.91	0.92	0.02	0.30
Activity Indicators:				
Articulation	-0.61	-0.50	0.40	0.23
Presentation	0.10	(0.60)	-0.97	(-.98)
Repetition	-0.72	-0.62	0.41	0.23
Verification	-0.70	-0.42	0.61	0.50
Didactic Actions	0.70	(0.73)	0.72	(0.92)
Speech thru Speech	-0.72	-0.65	0.40	0.20
Action and Speech	0.07	0.48	0.989	0.987
Variety Indicators:				
Articulation	-0.54	-0.61	-0.75	-0.90
Presentation	0.81	(0.55)	0.23	(0.99)
Repetition	-0.50	-0.30	-0.42	-0.90
Verification	0.11	0.30	-0.41	0.40
Didactic Actions	0.80	(0.50)	-0.25	(0.998)
Speech thru Speech	0.10	0.20	-0.33	-0.63
Action and Speech	0.03	0.28	0.94	0.945

Group One (A, B, C, D, E) **Group Two (B, C, D, E)**

TABLE 27

Summary Table of Correlations

TEST Language Use A(i),(ii) Mother Tongue B(i), (ii)		LISTENING		SPEAKING	
		GROUP ONE	GROUP TWO	GROUP ONE	GROUP TWO
A (i)	Occurrence of Language Use (per minute of allotted time)	-0.40	0.05	0.90	0.83
A (ii)	Types oc language changed per minute of allotted time.	-0.30	0.06	0.70	0.61
B (i)	Occurrence of L1 per minute of allotted time		0.74		0.84
B (ii)	Proportion of L1 to total frequency of language used		0.74		0.82

Group One (A, B, C, D, E) **Group Two (B, C, D, E)**

Actions (didactic) had been very active in Lessons C, A, D and E; but showed great variety in Lessons A, C, E and D.

Speech thru speech showed the greatest activity in Lessons E, C, B, A and D and the greatest variety in Lessons E, C, B, A and D.

Action and speech showed the greatest activity in Lessons B, C, E, A and D and the highest variety in Lessons C, B, E, A and D.

Therefore, there were two behaviours active and four varied in Lesson A; whereas, in Lesson B all five behaviours were comparatively active, but only two of them were varied. And six behaviours were active and varied out of the seven identified in Lesson C. Whereas in Lesson D two behaviours were active out of the seven identified, and four were varied. But, in Lesson E there were six very active behaviours and five of them showed great variety.

3.1 *Correlation of Behaviour and Mean Gain Scores*

Correlation coefficients had been calculated between the behaviour indicators of the lessons and the mean gain scores on the listening comprehension and the speaking tests. These correlations were made as an ultimate test of the hypothesis which stated, that the degree of linguistic performance of 11 year olds, in a beginning class of English as a second language, was due to the effects of didactic behaviours during the teaching/learning processes.

Therefore, Pearson's Product-Moment formula was used in the claculation of these correlation coefficients. The results of this test showed that very active and intense lessons might have brought about the desired change in learners' linguistic performance. And while a lesson which was greatly varied might have had good effects on the speaking test, it did not have the same results on the listening comprehension test in Group Two, but a less varied lesson might have been effective in

Group One. And the high negative correlations showed that frequent mistakes during learning might have caused learners to score less on the speaking test. Similarly, frequent errors might have caused increased errors on the listening test in Group Two, while less mistakes might have caused less errors in Group One. And the high positive coefficients indicate that a high percentage of errors made during learning might have caused learners' utterances to contain a high percentage of errors of vocabulary and grammar.

Correlation coefficients had also been found for the behaviour indicators of the specific behaviours and the tests. They are as follows: Articulation which was very active might have been effective on the speaking test, but too much variety did not have the same effect. Similarly, too much activity and variety during articulation did not bring about the desired change in learners' linguistic performance (listening comprehension).

Presentation that was too active might not have brought about the expected results on the speaking test; the same interpretation applies to Group Two on the listening test. But, in Group One, it seems that a less active presentation caused equally poor results on the listening test. However, it appeared that presentation which was varied might have been effective on the listening test, and the surprisingly higher correlations show that variety had been effective on the speaking test.

Repetition which showed too much activity and variety might not have been effective on the listening test. And whereas active repetition had been effective on the speaking test, too much variety might have caused adverse effects.

Verification which was active did not seem to have good effects, but much variety might have had good effects on the listening test. And while an active and varied verification might have been effective on the speaking test in Group Two, in Group One much activity was effective, whereas too much variety was not effective.

Didactic actions which were highly active and varied appeared to have been effective on the listening and speaking test. However, less variety in Group One might have caused similar effects on the speaking performance of learners in this group.

Speech thru speech which was very active but not varied, appeared to have been ineffective on the listening comprehension test. While much activity during drills had been effective, on the speaking test, very varied drills were not.

Action and speech that were active and varied had been effective in Group Two on the listening test, but less activity and variety in Group One might not have brought about the desired change in learners' linguistic performance on this test. And, the surprisingly high and positive correlations seem to indicate, that action and speech activities during learning, might have been effective in bringing about the desired change in the linguistic performance (speaking test) of learners in both Groups One and Two.

In addition, correlation coefficients had been calculated for teachers' language behaviour. The correlations were made between A; language use and the tests, B; the use of the mother tongue and the tests. The correlations show that very frequent use of language, and frequent changes of categories, might not have been effective on the listening test. However, the high occurrences of language use and change of categories, appeared to have been effective on the speaking test. Similarly, it appears that learners' contact with the mother tongue, which was frequently used during teaching/learning, might have brought about the desired changes in their linguistic performance. And, the interpretation above might apply when a high percentage of language, when that language was the mother tongue, was used in the teaching and learning processes.

3.2 *Findings and Conclusions*

The study revealed that learners' linguistic performance had been caused by the effective use of didactic behaviours, which were in terms of varying degrees of duration and frequency, during teaching and learning processes.

The pre-test revealed that the new didactic points were unknown to pupils. And while they showed some language skills on the listening test, learners obtained zero scores on the speaking test, with the exception of four pupils in Lesson C who scored a total of seven points, and five pupils in Lesson E whose total score was nine points. The post test data revealed that learners' speaking skills improved significantly, and all five classrooms appeared to have acquired equal language ability. However, the test data revealed that wide differences existed in the development of listening comprehension skills, among learners of the five classrooms; or that teachers had used procedures which might have been different.

The analyses of the five classrooms revealed two important differences: 1. The study showed that teachers A, C, D and E made use of seven behaviours, and the new materials were introduced during the time-slot 'presentation'. While teacher B used only five behaviours and introduced the new didactic points during 'repetition'. 2. Teachers B, C, D and E taught all the new vocabulary and grammar, whereas teacher A presented the two grammar words only. In addition, teacher A used the target language only during teaching, while teachers B, C, D and E used the mother tongue as well. And whereas, it was restricted to praise or disapproval by teachers C, D and E, L_1 was the medium of communication used by teacher B. Moreover, teacher B's comment *"...pas un mot en français!"* to pupils who spoke French, appeared ironical.

The average or mean score was the unit of measure used throughout the study project. Therefore, test scores, behaviour indicators and correlations are averages which have different functions in this study. Moreover, during the discussion of the findings, lessons will be placed in positions which are determined by the highest numerical value of their averages.

Learners' linguistic performance on the comprehension test revealed that Lesson A > B > C > D > E >, in Group One, while Lesson B > C > D > E in Group Two. In Group One, learners' mean gain score had been the highest in Lesson A, and in Group Two, the mean gain score of greatest value was Lesson B.

In addition, learners' performance on the speaking test revealed that Lesson B > C > E > A > D, in Group One, and Lesson B > C > E > D, in Group Two. Therefore, learners in Lesson B showed consistency in obtaining the highest successes on the speaking test, for both groups.

The study revealed that highly active lessons might have brought about the desired change in learners' linguistic performance (listening/speaking). Thus, activity in Lesson B > A > C > D; therefore, Lesson B gave the highest evidence of activity in support of the hypothesis, while Lessons A, C, E and D, were less active but they also supported the hypothesis.

Likewise, lessons that were very intense caused the desired change in learners' linguistic performance (listening/speaking). The density showed in the use of the new didactic points were greatest in Lesson B > A > E > C > D. The close contact of learners, with the new didactic materials in action and speech activities, during learning, in Lesson B, caused a higher incidence of success on their linguistic performance, than the evidence shown in the other lessons.

It was found that a highly varied lesson might have brought about the desired results on the speaking test, but not on the listening test of Group Two. And less variety had produced good results in group One. Hence variety in Lesson C > E > B > A > D. However, variety had been almost equal in all five lessons, but there were marginal differences. Therefore, Lessons D, A and B gave evidence of less varied lessons, but all lessons showed, in general, a lack of variety. And Lessons C and E gave evidence of most variety in support of the hypothesis.

The study also revealed that individual learners' performance was closely related to their linguistic performance. It was found that correct utterances which occurred often during learning were closely related to the superior performance on listening and speaking tests. Hence, Lesson B > C > E > A > D. Therefore learners in Lesson B support the hypothesis to a greater extent than learners in Lessons C, E, A and D. Also when a high percentage of individual utterances was correct during learning, this also had good effects on linguistic performance; this ratio in Lesson E > C > A > B > D. Therefore, learners in Lesson E gave the highest evidence in support of the hypothesis, and learners in Lesson C, A, B and D did so, but to a lesser degree. The study also showed that a high percentage of mistakes made during learning, was closely related to similar mistakes on the listening comprehension test, but, there were fewer errors on the speaking test; the ratio in Lesson D > B > A > C > E. Therefore, learners in Lesson D and B made less errors on the speaking test but more on the listening comprehension test, and this was the reverse in Lessons E, C and A. However, a high percentage of error of vocabulary and grammar during learning, appeared closely related to correspondingly high incidence of these errors on the listening and speaking tests; the ratio of incorrect vocabulary and grammar in Lesson A > D > C > B > E. Since learners in Lessons E, B and C had the lowest proportion of mistakes, the hypothesis would be supported in these lessons.

The effects of activity and variety of the specific behaviours on learners' linguistic performance will be discussed in the following paragraphs.

The study showed that articulation which was active, was related to success on the speaking test, but not on the listening test. And an active articulation in Lesson E > B > C > A > D. Therefore, the hypothesis was supported in Lesson E which showed the highest activity, and in Lesson D in which there had been least activity during articulation. However, too much variety did not bring about the desired change in learners' linguistic performance (listening/speaking). And variety in Lesson D > E > A > C > B; hence, the hypothesis was supported in Lesson B which showed the least variety, and in the four other lessons as well, but to a lesser degree.

The study also showed that an active presentation was effective on the listening comprehension, but not on the speaking test. And activity in Lesson A > D > E > C. The hypothesis is supported in Lesson A and also in Lesson C. And variety during presentation was closely related to learners' successful linguistic performance on both tests. The variety in Lesson A > C > D > E; therefore, the hypothesis was supported in the highest degree in Lesson A and to a lesser degree in Lessons C, D and E.

The study showed that active repetition was effective on the speaking but not on the listening tests. Active repetition in Lesson E > C > B > A > D; therefore, the hypothesis is supported in Lesson E which showed the highest activity and also in Lesson D, in which there was less activity. However, the study also revealed that a very varied repetition might not have brought about the desired change in learners' linguistic performance. And variety during repetition in Lesson D > E > C > B > A. Therefore, the hypothesis is supported in Lesson A which showed the least variety during repetition, but Lessons B and C also demonstrated less variety.

The study also revealed that activity during verification was effective on the speaking test but not on the listening test. And active verification in Lesson E > C > B > D > A. Therefore the hypothesis is supported in Lesson E, since this lesson shows evidence of greater activity, and Lesson A, and to a lesser degree D and B, also support the hypothesis. It was also found that variety during verification was effective on the listening test and on the speaking test of Group Two but not in Group One. Variety during verification in Lesson C > D > A > E > B; therefore, the hypothesis is supported in Lesson C and also in Lesson B, and to some extent in the other lessons.

The study also revealed the activity during didactic actions had been effective in bringing about the desired change in learners' linguistic performance (listening/speaking). Activity of didactic actions in Lessons C > A > D > E. The hypothesis is supported in Lesson C which gives evidence of the highest activity, as well as in the other lessons to a lesser degree. And didactic actions which were varied also had the same effects on learners' linguistic performance, except for speaking in Group One. Varied didactic actions in Lesson A > C > E > D. Therefore, the hypothesis is supported in Lesson A and also in Lesson D.

The study revealed that active speech thru speech (drills) had been effective on the speaking but not on the listening test. And active drills in Lesson E > C > B > A > D. Therefore, the hypothesis is supported in Lesson E which showed evidence of highest activity and in Lesson D, with the least active drills. But, it was found that variety during drills was not effective on the speaking test, and less variety was not effective on the listening test. Variety during speech thru speech in Lesson D > C > B > E > A. And while much variety might have been effective on the listening, test less might predict good results on the speaking test. Therefore, the hypothesis is supported in Lesson D and also in Lesson A.

The study revealed that action-and-speech which were active, might have brought about the desired change in learners' linguistic performance (listening/speaking). Activity during action-and-speech in Lesson B > E > A > D > C; supports, the hypothesis in Lesson B to a higher degree than in the other lessons. Variety during action-and-speech was very effective; this was particularly so in Lesson C > B > E > A > D. Therefore, the hypothesis is supported in Lesson C more than in the other lessons.

It was also found that teachers' language use, the purpose for which language was used, had been effective in bringing about changes in learners' linguistic performance. Therefore, teachers' frequent use of language, every minute, and, also frequent changes of categories, had been effective on the speaking but not on the listening test; and the highest indication of this is shown in Lesson C > B > E > D > A. Therefore, Lesson C and Lesson A support the hypothesis. The study also showed that frequent use of the mother tongue had been effective in bringing about changes in linguistic performance; the greatest effects were observed in Lesson B > C > E > D. Therefore, the hypothesis is supported in Lesson B to a higher degree than in the other lessons. Likewise, it was found that the close contact of learners with L_1, when it was used in a high proportion during teaching and learning, had been very effective, in Lesson B > C > D > E; once more the hypothesis is supported in Lesson B.

In conclusion, the evidence shown above, for the five ESL classrooms, supports the hypothesis, that learners' linguistic performance had been caused by the effective use of didactic behaviours, and also teachers' language behaviour, the use of the mother tongue more specifically. However, a brief survey of the study will help to highlight the lesson which demonstrated maximum support of the hypothesis.

The hypothesis was supported by the greatest number of observable facts in Lesson B. In this lesson evidence of effectiveness was given nine times on the listening test and ten times on the speaking test. Whereas, for Lesson A, evidence was given six times on the listening and twice on the speaking test. And for Lesson C, evidence of effectiveness was given three times on the listening and five times on the speaking test. In Lesson D, the hypothesis was supported four times on the listening and twice on the speaking test. And, there were three supporting evidence on the listening and six on the speaking test, in Lesson E. It follows then, that the maximum evidence of effectiveness of didactic behaviours and language use (L_1), was shown in Lesson B, where, for example, B > A > D > C > E, on the listening test and B > E > C > A > D, on the speaking test. This is further substantiated by the fact that learners in Lesson B obtained the highest gain scores on speaking, as a result of their close contact with the new didactic points during teaching and learning; this was true for Group One and Group Two. On the other hand, the gain in the successful recognition of the new materials was greater in Lesson A than in B, in Group One, but learners in Lesson B scored highest in Group Two. However, the five ESL classrooms gave sufficient evidence in support of the hypothesis. The study showed that the functional use of behaviours during the teaching/learning processes, had been effective in bringing about the desired change in learners' language performance.

Yet it still remains to be seen which behaviours, how many and how much of each are most effective. Do some theoretically good lessons produce more learning, or is it a question of some types of teaching behaviour yielding more learning than other types? This will be the question of our next chapter.

Chapter 5

DO SOME TYPES OF TEACHING BEHAVIOUR YIELD MORE LEARNING?
Evidence from Polychronometric Experiments

1. **The Learner's Language Use**

2. **The Teacher's Language Behaviour**

3. **Error Correction**

4. **Variety of Activities**
   4.1 *Third-Grade Learners*
   4.2 *Ninth-Grade Learners*

Teachers teach on the assumption that learners learn what they have been taught. In foreign language teaching this has been a grand illusion. So few learners actually learn a foreign language during the time they spend in class that one wonders whether the returns are worth the investment. Yet the few exceptional classes which produce some fluent speakers of a foreign language lead us to ask why such classes have to be so exceptional. Is it the method, the techniques, the personality of the teacher, some or all of these, and in what proportion? In other words, why do some classes of learners actually succeed in learning a foreign language in class. Is it because they have given their attention to few of the most important structures and words and learned them well? Is it because they have mastered the system of the new language in such an order as to produce the maximum capacity of expression with the minimum amount of interference? Is it because they were presented the language in such a way as to remember it well? Is it because they made maximum immediate use of what they had been taught? Or is it simply because they learned the language from a "good" teacher?

It has been said that no method can replace a good teacher. A teacher's good judgment based on long and successful experience with a particular type of class is often more reliable than are the doctrines of applied linguistics and teaching psychology. This having been said, the burden of proof still remains when opinions differ on how languages should be taught; and when differences in teaching practices may be matters of degree, there remains the problem of measurement. This is the business of research.

In the field of language didactics, research has been based on the analysis of the input and of the corresponding output of formal language teaching and learning, as distinct from the observation of natural language learning such as one finds, for example, in a bilingual family. The input of the formal learning process comprises the materials used by the learner and also what the learner sees and hears from the teacher. Both types of input can vary in several respects to produce quite different results. The observable output of the language learning process consists of the oral and written utterances of the learner and the progressive modification of which is supposed to end up in a fluent, correct and independent use of the new language based on a clear comprehension of what has been heard and read. Auditory or written comprehension may, however, be the only objectives, in which case the analysis need not go beyond the emission-reception relationship observable through various types of feedback control. The objectives, whatever they may include, are generally to be found in some sort of syllabus, method or teaching manual. Theorists, such as Bialystock (1978), Krashen (1981) and Hammerly (1985), have constructed theoretical models to describe the internal learning process. Others, such as Rivers (1972), Breen and Candlin (1981), and Stern (1983), have proposed instructional models to describe effective classroom processes. While theoretical positions are valuable as potential descriptions of the teaching-learning process, experimental research is necessary for theory verification.

Investigations into the second language classroom interaction process have normally been general in nature, with no attempts being made to analyze individual lessons as a means of determining specifically the reasons for behaviour changes. The four short-term experimental studies discussed in this chapter, however, are based on detailed polychronometric analyses of 40 lessons in order to identify possible relationships between classroom behaviours and measured learning. These investigations involved a pretest-posttest procedure to measure the learnings of third and ninth grade subjects, with psychometric data submitted to an analysis of co-variance where the co-variates were sex, age, IQ score, attitude score, and father's occupational classification. The classroom behaviour quantifications were subsequently used to interpret the statistical data.

1. The Learner's Language Use

In one study, matched groups of beginning second language students participated in lessons in which learners' use of the target language was extensive (Lesson A1) or restricted (Lesson A2): Analysis of covariance conducted on test scores indicated a tendency for subjects who participated in lessons where the target language was most frequently employed (Lesson A1) to have superior test scores. For oral comprehension we found p. $< .11$ and for oral production we obtained $p < .05$, both favoring learners who had been involved in the A1 lessons.

As a means of meaningfully interpreting these findings, we quantified various categories of classroom behaviours. In order to provide a broad picture of our lessons and permit comparisons of the distribution of classroom behaviours, we first of all analyzed general didactic behaviours. Our analysis did reveal significant behavioural differences between the two lesson types, A1 and A2. In both types of lessons we normally found important differences in the distribution of the teaching categories known as Semantic Presentation (SP), Formal Presentation (FP), and Semantic Repetition (SR), and of the learning categories identified as Learner Action Responses (AC), Learner Speech Responses (SH) and Choral Repetitions (CH), as shown in Table 28.

TABLE 28

Total Distribution of General Didactic Behaviours in Lessons A1 and A2

Behaviour Categories	GENERAL DIDACTIC BEHAVIOURS HAVING MAJOR DISTRIBUTION DIFFERENCES			
	Lesson A1		Lesson A2	
	Duration	Frequency	Duration	Frequency
Teaching				
SP	146.2	92	182.8	109
FP	161.7	95	218.5	120
SR	88.1	76	155.8	120
Learning				
AC	392.8	76	539.9	91
SH	747.1	557	327.6	253
CH	395.3	226	101.1	84

As Table 28 shows, the distribution of teacher didactic behaviours and learner action responses was greater in lessons known as A2, while that of learner verbal didactic behaviours was higher in lessons referred to as A1.

If instructor target language usage had been the major factor influencing learning, we might have expected presentation behaviours to be most influential. In our A1 lessons, presentation behaviours occurred a total of 187 times and had a total duration of 307.9 seconds; in the A2 lessons, they appeared a total of 229 times and lasted a total of 401.3 seconds. Although the distribution of teacher presentation behaviours was greater in the A2 lessons, learning was normally superior in the A1 lessons, leading one to feel that teacher presentation behaviours might not have been the most important determiners of learning.

Besides general presentation behaviours, the distribution of most presentation sub-categories was also greater in the A2 lessons (see Table 28).

As we see in Table 29, except for PT behaviours, all types of presentation behaviours generally occurred more frequently and lasted longer during the A2 lessons.

TABLE 29

Total Distribution of Presentation Behaviours in Lessons A1 and A2

Behaviour Categories	DISTRIBUTION OF PRESENTATION BEHAVIOURS			
	Lesson A1		Lesson A2	
	Duration	Frequency	Duration	Frequency
PA	121.1	74	189.3	79
PT	60.7	52	60.5	43
PX	23.4	22	40.5	30
PO	33.8	30	47.7	39
OM	45.1	31	53.1	39

From the information presented thus far in our discussion, it seems as if differences in the distribution of instructor presentation behaviours did not, by themselves, account for test score variations. If they had, we would have anticipated superior test performances from subjects who had participated in the A2 lessons.

Besides the presentation phase, an instructor may utilize the target language during other phases of a lesson. As we saw earlier, when we analyzed and quantified general didactic behaviours, the teacher employed SR behaviours more frequently in the A2 lessons. However, after we sub-categorized and analyzed all teacher second-language utterances we found no overall differences which we felt might have been related to variations in learning.

While differences in the distribution of teacher didactic behaviours did not, at least in themselves, seem to play a major role in influencing test score variations, it is possible that dissimilarity of instructor language behaviours may have been, at least partly, responsible. As we see in Table 30, however, few major behavioural differences existed between the two lesson types.

In Table 30 we see that, except for frequency of MT behaviours, only minor differences existed between the distribution of instructor language behaviours in the two lessons. Since much of the difference between the frequencies of mother tongue usage can be accounted for by examining articulation behaviours, one might not have anticipated that these additional occurrences of the MT category would have been the major determiners of learning differences.

TABLE 30

Total Distribution of Instructor Language Use Behaviours in Lessons A1 and A2

Behaviour Categories	INSTRUCTOR LANGUAGE USE BEHAVIOURS			
	Lesson A1		Lesson A2	
	Duration	Frequency	Duration	Frequency
DE	231.3	162	311.0	175
MT	232.0	150	222.4	114
QL	108.4	112	121.8	117
IT	316.5	274	344.8	269
CP	25.0	23	9.8	8

Our analysis of instructor behaviours seemed to indicate that they, at least by themselves, did not account for test score variations. However, in a formal classroom setting, learners, as well as instructor, normally participate in the interaction process. That being the case, variations in learner classroom behaviours might have significantly influenced test scores.

During each lesson, all subjects involved in the current investigation had occasion to respond verbally and/or physically to certain stimuli. As we see in Table 31, large differences existed between the total distributions of certain behaviour categories.

TABLE 31

Total Distribution of Learner Actions and Speech Behaviours in Lessons A1 and A2

Behaviour Categories	LEARNER ACTIONS AND SPEECH			
	Lesson A1		Lesson A2	
	Duration	Frequency	Duration	Frequency
KS	524.2	407	230.6	175
LD	392.9	76	540.0	91
LR	102.2	95	43.7	36
RT	386.0	219	88.5	75

As Table 31 indicates, subjects involved in Lesson A1 tended to have more verbal responses, while those participating on Lesson A2 normally had a higher distribution of physical behaviours.

Subjects involved in Lesson A1 not only had a higher distribution of target language behaviours, but they also had a greater distribution of each behaviour type. In Table IV we saw that each verbal behaviour category occurred more often and lasted longer in Lesson A1.

In our study, the learner verbal activity involving the greatest amount of semantic content was the one classified as KS, *Look and Say*. Subjects involved in A1 lessons generated 232 more occurrences of that behaviour category than those in A2 lessons, and its total duration was 293.6 seconds longer in the same lessons. This type of learner didactic activity may have helped subjects internalize the semantic content of the teaching-points, in addition to providing meaningful speaking practice.

The subjects participating in Lesson A1 seem to have been more frequently, and for a longer period of time, overly involved with meaningful content than those participating in Lesson A2.

Since most learner second-language utterances occurred in the presence of referential materials, even activities whose primary goal was formal practice may have been accompanied by a considerable amount of semantic content. If that is the case, subjects who found themselves involved in Lesson A1, may have been exposed to a significantly greater amount of semantic content than those in Lesson A2. Considering the views that meaningful content is learned faster and retained longer than less significant material, and that active practice is superior to passive practice, the learners' active participation in highly meaningful activities might have affected learning. This type of didactic behaviour might have been effective because it required learners to comprehend content and mentally process information in order to provide correct utterances.

It may be possible that, as Krashen (1980) suggested, where total language input is high, subjects learn the target language faster. However, it seems as if learning may be greater where learners provide a substantial amount of that input. If this were not the case, we might possibly have anticipated superior test performances form learners involved in Lesson A2, where the instructor supplied a higher percentage of the second-language input.

Learner language behaviours, as well as didactic behaviours, may also influence learning. Our quantification of student language behaviours indicated a number of marked behavioural differences (see Table 32).

TABLE 32

Total Distribution of Language Performance in Lessons A1 and A2

Behaviour Categories	LANGUAGE PERFORMANCE BEHAVIOURS			
	Lesson A1		Lesson A2	
	Duration	Frequency	Duration	Frequency
PE	124.1	93	56.4	38
VE	24.0	11	5.4	4
OK	500.5	400	241.2	187
CC	32.4	31	17.3	15
SI	4.2	3	0.0	0

Since, for this level of analysis, we did not consider phonetic inaccuracies to be miscues, subjects involved in Lesson A1 produced a total of 527 acceptable individual target language utterances, while those in Lesson A2 provided 240, for a difference of 287. Not only did subjects participating in Lesson A1 produce more second-language utterances than those in Lesson A2, they also generated many more correct ones.

The information gathered from our lesson analysis seems to indicate that the most influencial factor affecting test performances was the learners' successful active involvement in meaningful activities. In Lesson A2, although the instructor spent more time presenting and practicing the teaching-point and learners had more physical responses, subjects generally performed less well on tests. In groups where learners asked and correctly answered many questions in meaningful contexts, responded to more teacher questions, and provided speech utterances to accompany the physical behaviours of classmates, test scores tended to be higher on both aural comprehension and oral production tests, significantly so on the latter.

While subjects having more opportunities to utilize the target language tended to perform better on both types of tests, the amount of speaking, per se, might not have been the crucial factor deciding how well learners acquire items within a syntactic structure. The type of language they employ, which may be related to the degree of active learner involvement, could be an important contributing factor. Where the distribution of learner speech was higher (and test performance superior), subjects also participated in more semantically-oriented activities.

2. The Teacher's Language Behaviour

A second study was designed to examine the effects of instructor code-switching on student learning. In this investigation matched groups participated in lessons in which the instructor employed only the target language (lesson F3) or utilized both the learners' native and second languages (lesson F4). Analyses of co-variance yielded $p < .001$ for listening comprehension, favoring subjects exposed to instructor code-switching behaviours; for oral production, differences were negligible ($p < 0.97$). Classroom behaviours were subsequently quantified to determine possible explanations for these findings.

Our initial analysis of general didactic behaviours revealed substantial behavioural differences (see Table 33).

TABLE 33

Distribution of Certain General Didactic Behaviours in Lessons F3 and F4

Behaviour Categories	GENERAL DIDACTIC BEHAVIOURS			
	Lesson F3		Lesson F4	
	Duration	Frequency	Duration	Frequency
Teaching				
SP	169.1	80	307.9	162
FP	194.1	86	186.4	111
FR	475.2	238	477.4	317
Learning				
SH	483.0	341	510.8	360
CH	386.0	210	270.8	157

As we see in Table 33, except for the CH classification, the distribution of all behaviour categories showing widely different quantifications was higher during F4 lessons.

Our analysis of more comprehensive didactic behaviours provided useful information and seemed to indicate that in the F4 lessons total input might have been greater then in the F3 lessons. To obtain more specific data, however, we undertook finer categorizations, with subsequent quantification.

Since one of the primary goals the instructor is presentation is to get content into the heads of learners, and since subjects involved in Lesson F4 outperformed those who participated in Lesson F3 on aural comprehension tests, our initial second-order analysis was of teacher presentation. While the distribution of most presentation behaviours was similar, we did observe major distribution differences with PA and PT categories (see Table 39).

TABLE 34

Distribution of Presentation Behaviours PA and PT during Lessons F3 and F4

Behaviour Categories	PRESENTATION BEHAVIOURS			
	Lesson F3		Lesson F4	
	Duration	Frequency	Duration	Frequency
PA	244.0	81	186.2	79
PT	0.0	0	121.5	117

In Table 34, we observe that PA behaviours lasted 57.8 seconds longer in Lesson F3 than in Lesson F4, while PT behaviours were frequent in Lesson F4 but not-existent in Lesson F3.

As we indicated above, during F3 lessons, the instructor spent a considerable amount of time performing various actions in order to get meaning across to the learners. This was necessitated because of the requirement to employ only the target language to present content. In Lesson F4, however, where SP activities were more frequent and longer-lasting, and the instructor freely utilized the learners' mother tongue, it was not necessary to spend as much time employing PA behaviours in order to present the content.

In F4 lessons the instructor employed the learners' native language 117 times with a total duration of 121.5 seconds during presentation activities. This meant that from the outset of the lesson, learners comprehended what they were expected to learn, with little guessing involved. In the F3 lessons, however, learners were constantly conjecturing, often incorrectly, about the content being taught. In lessons where the instructor utilized both the target and native languages, learner hypotheses were constantly being confirmed, or rejected and subsequently revised. As a consequence, both instructor and learner utterances, which were often accompanied by use of pictures or objects, carried a great deal of known semantic content. For subjects involved in Lesson F3, however, much of the classroom repetition was virtually meaningless because they had problems determining exactly what was being said. Possibly, because it enabled subsequent utterances to carry much semantic content, instructor use of learners' mother tongue during presentation might have significantly influenced listening comprehension scores.

After further second-order quantifications of teacher didactic behaviours failed to provide other useful information that might supply explanations for mean test variations, we analyzed instructor language behaviours. Again, the only significant language difference between the two lesson types was the distribution of MT behaviours, which occurred 291 times and lasted 454.4 seconds in Lesson F4 but only twice with a duration of 0.9 seconds in Lesson F3. The distribution of teacher language behaviours seemed to verify that the greatest difference between the two lessons, from the instructor's perspective, was the language employed in the classroom.

If instructor behaviours were major determiners of listening comprehension, utilization of learners' mother tongue seems to have had the most significant influence. Apart from this, teacher behaviours in both types of lessons generally seemed to be similar.

In a typical classroom setting, however, learners, as well as teachers, exhibit behaviours which may influcence learning. Since our earlier analysis of learner general didactic behaviours indicated behavioural variations between lessons, we further categorized and quantified learner behaviours. This analysis of didactic behaviours indicated that while most categories were distributed similarly in the two lessons, RT and TR behaviours showed considerable variation (see Table 35).

TABLE 35

Total Distribution of Learner RT and TR Behaviours in Lessons F3 and F4

Behaviour Categories	LEARNER SPEECH BEHAVIOURS			
	Lesson F3		Lesson F4	
	Duration	Frequency	Duration	Frequency
RT	386.0	210	251.8	143
TR	0.0	0	35.7	27

In Table 35 we see that RT behaviours were more frequent and longer-lasting in F3 lessons, while TR behaviours occurred 27 times, for a total duration of 35.7 seconds in F4 lessons, but were non-existant in the other lessons.

The distribution of RT behaviours might have influenced listening comprehension scores. In lessons where the instructor had employed the learners' mother tongue (F4 lessons), especially during presentation, any subsequent student utterances would probably have been meaningful and, as such, might have positively influenced learning. In the F3 lessons, however, although subjects might have utilized the target language more frequently, the utterances might not have induced as much learning because they failed to carry sufficient semantic content to facilitate internalization. (In fact, aside from RT behaviours, the distribution of meaningful learner activities was higher in Lesson F3, with 1012 occurences, as opposed to 956 in Lesson F4). In this way, it seems as if instructor use of learners' native language might have further affected learner comprehension.

Another learner behaviour which seems to have influenced comprehension was the one categorized as TR. Employment of this behaviour type by learners was generally followed immediately by confirmation or correction. As such, it frequently enabled subjects to determine immediately what the appropriate meaning of an utterance was in their native tongue.

With learner didactic behaviours, due to variations in the distribution of instructor usage of learners' native language, apparently having played some role in determining mean listening comprehension differences, we attempted to determine the effects of student language performance on learning. Quantification of various learner language behaviour categories indicated, however, that only minor differences existed between groups involved in each type of lesson. The magnitude of differences was such that, we felt, it would not have been a major influence on variations in levels of learner listening comprehension, but might have been related to the overall similarity of mean oral production scores.

With few major learner behavioural differences occurring between the two lesson types, one might have felt that learner actions and speech had played little, if any, role in producing learning differences. However, because of prior mother tongue usage by the instructor in Lesson F4, every learner utterance in these lessons probably carried considerably more semantic content than in the F3 lessons and might have in this way, increased aural comprehension skills. As well, learner TR behaviours might have further increased student understanding of semantic content. However, learner language performance did not seem to have seriously affected learning.

The primary source of variation in subjects' level of listening comprehension appears to have been the distribution of the language(s) employed by the teacher during the language lesson. The mother tongue usage may have created conditions which made total input, semantic, input and learner repetitions, more meaningful for these students. From the early part of the lesson, subjects involved in Lesson F4 understood the meaning of what they were hearing and practising, and when they arrived at the end of the lesson they had heard and practised the already-understood content a sufficient number of times to perform better on the test. When the test came, students were not disadvantaged by having to respond quickly to an utterance whose meaning they had incorrectly surmised or had acquired only a couple of seconds earlier in the actual test situation.

The effects on aural comprehension of presence or absence of learners' mother tongue during a language lesson seemed to vary depending on the chronological age level of subjects. Possibly because third grade learners are still functioning at the concrete operations stage of mental reasoning, learners involved in the lesson in which their first language was not employed, appeared to have had difficulty extracting meaning, especially of the verb *faire*, from actions or pictures. From the researcher's observations of student verbal behaviours during the lessons, it was apparent that many third grade learners felt uneasy, not being certain of the meaning of words, utterances, or phrases which they or the instructor employed.

The older subjects, operating at a more advanced stage of mental operations, seemed to have experienced little difficulty comprehending the teaching-point in the absence of their mother tongue. Inversely, its utilization did not seem to help learners gain higher test scores, with both groups obtaining identical mean test scores in aural comprehension.

3. Error Correction

In a third investigation matched groups of learners participated in lessons in which the instructor provided either immediate (V5 lessons) or delayed (V6 lessons) error correction. Analysis of co-variance yielded $p <.73$ and $p <.05$ for aural comprehension and oral production respectively, both favoring students involved in lessons in which corrective treatment was immediate (V5 lessons). Classroom behaviours were subsequently quantified to find possible explanations for these results.

Initially, presentation behaviours were examined to identify potential causes for group similarities in mean aural comprehension test scores. When we totaled the figures representing the duration and frequency of presentation behaviours, we found that semantic presentation activities occurred 81 times and lasted a total of 156.3 seconds, and formal presentation behaviours 99 times and lasted a total of 206.4 seconds, for a grand total of 180 occurrences lasting 362.7 seconds in V5 lessons. In the four V6 lessons, having later correction, semantic presentation behaviours occurred 70 times, lasting 134.4 seconds, and formal presentation behaviours 77 times, lasting 159.9 seconds, for a grand total of 147 presentation behaviours having a total duration of 294.3 seconds. From these figures we see that both categories of presentation behaviours had higher distributions in the lessons having immediate error correction.

To obtain more specific information we further categorized and analyzed teacher presentation behaviours. This analysis indicated that most behaviour categories had higher distributions in the V5 lessons, some significantly so (see Table 36).

As Table 41 indicates, behaviour categories exhibiting wide variations had higher distributions in the V5 lessons. If the distribution of presentation behaviours alone had determined test scores, we would probably have expected superior performances form subjects involved in lessons where the instructor employed immediate error correction behaviours.

While the distribution of presentation behaviours may affect learning, other teacher didactic behaviours could also play a role. With this in mind, we quantified various categories of other teacher repetitions, some of these, as well, showing widely different distributions. In total, teacher repetitions occurred 436 times, lasting 691.2 seconds in lessons classified as V6; in Lesson's V5 we found 369 incidents of similar behaviours, lasting a total of 604.1 seconds. As the figures indicate, the distribution of specific instructor repetition behaviour categories was usually higher in the V6 lessons.

TABLE 36

Total Distribution of PX, PT and OM Behaviours in Lessons V5 and V6

Behaviour Categories	SPECIFIC PRESENTATION BEHAVIOURS			
	Lesson V5		Lesson V6	
	Duration	Frequency	Duration	Frequency
PX	72.3	49	53.4	37
PT	56.2	40	39.5	27
OM	66.2	42	51.3	33

After quantifying presentation and repetition behaviours, we analyzed instructor utterances which Mackey (1978) classified as verification. During the lessons under consideration, ten such behaviour categories occurred, but only Prompting (PR), Remodelling (RM) and Confirming (CO) showed much distributional variation (see Table 37).

As Table 42 indicates, the three categories having considerable variation in distribution were more frequent and longer-lasting in the V5 lessons.

After totalling instructor didactic behaviours and finding that the total distribution was similar, we categorized and quantified teacher language behaviours. This analysis confirmed the overall distributional similarity observed in the quantification of didactic behaviours.

TABLE 37

Total Distribution of PR, RM and CO Behaviours During Lessons V5 and V6

Behaviour Categories	SPECIFIC VERIFICATION BEHAVIOURS			
	Lesson V5		Lesson V6	
	Duration	Frequency	Duration	Frequency
PR	21.1	11	7.3	6
RM	77.4	93	60.3	56
CO	137.3	205	108.7	164

Subsequently we analyzed and quantified student didactic behaviours, but found only one significant distributional variation. That difference occurred in the distribution of RT behaviours, a formal repetition category; they occurred 149 times in V5 lessons, and 122 times in V6 lessons. With the exception of RT, the total distribution of each category of action and speech behaviours in the two lesson types being considered, was similar. When we totalled the figures for each classification of behaviours in the two different types of lessons, we discovered that learners involved in lessons having immediate error correction used the second language 618 times, lasting 977.0 seconds. In the other lessons we found 579 occurences of student second language speech, having a total duration of 1058.4 seconds. As the figures indicate, student second language utterances were more frequent in lessons having immediate error correction but were longer-lasting in lessons having later error correction.

Subsequently we quantified the oral target language input to which subjects were exposed in both lesson types. In V5 lessons there were 1234 occurrences of the second language, lasting a total of 1981.1 seconds; in V6 lessons there were 1203 incidents, having a total duration of 2078.5 seconds. With such similarity in the distribution of total input, one would probably not have anticipated significantly different mean aural comprehension test scores.

Language lessons were also analyzed to account for the significant difference ($p < .05$) in mean oral production test scores. As we have already seen, students used the second language more frequently in lessons having immediate error correction, but the total duration of their utterances was longer in lessons having later error corrections. Overall, the greatest distribution difference occurred in the RT category. Largely because of differences in the distribution of this behaviour type we found that students in lessons known as V5 had a mean of 152.8 second language repetitions while those in V6 had a mean of 125.8 utterances. On the whole, students used the second language more often in lessons having immediate error correction. Although most of the mean distribution variation was due to differences in the duration and frequency of the RT category (behaviours which provided formal expression practice), most of these behaviours occurred in the presence of inferential materials. That being the case, each learner utterance could have carried considerable semantic content and positively influenced aural comprehension and oral production scores.

After analyzing learner didactic behaviours, we examined individual language performance. Behaviour quantifications indicated that the distribution of correct and incorrect individual second language utterances in each lesson type was nearly identical.

To this point we have attempted to explain differences in mean test score in terms of duration and frequency of behaviours without reference to error correction. It may be, however, that when instructors correct learner errors, they create conditions which facilitate or retard second language learning. Wirst of all, from the perspective of the teacher, if he or she interrupts students each time they make any type of error and provides corrections by means of emphasis, they are likely to make frequent, but short, second language utterances. Secondly, from the students' point of view, when learners are interrupted and corrected, they usually stop and repeat all or part of the utterance at least up to the point of correction. Sometimes within a single sentence a subject may make two or three miscues which, in immediate correction classes, are corrected as soon as they occur, leading the student to have a number of repetitions of all or part of the sentence. When errors are corrected after the student has completed the entire sentence, the teacher permits the learner to produce utterances at his own speed while possibly making several errors in the repetition; afterward he corrects all the errors, or as many as he remembers, at the same time. This may be why repetitions in lesson V6 tended to last longer than those in V5. By correcting students immediately the teacher created conditions whereby the students used the second language more frequently, but with the repetitions tending to be of shorter duration.

Another factor which came into play in this study was the percentage of errors corrected. Although we had instructed the teacher to correct all student errors and had provided practice sessions to improve correcting techniques, we found a considerable difference in the percentage of errors corrected during the two types of lessons. When providing immediate correction the instructor corrected 83 percent of student errors; when permitting students to finish their utterances, the teacher corrected only 51 percent of the errors committed. When learners make errors a teacher quickly decides whether or not to treat them. In lessons having later error correction, whether because of forgetting, concentrating less on detecting errors, or some other reason, the teacher corrected only about half of the errors. By doing this, he reduced considerably the number of second language utterances the student heard and produced.

While the delay between error commission and correction might not have directly determined student learning, it might have affected the amount of student input and output. Thus the effects of error treatment on learning seem to have been indirect. When error correction occurred immediately, subjects usually generated more second-language utterances. Immediate correction techniques created conditions whereby learners utilize the target language more often. Since most utterances were accompanied by pictures and/or actions, each speech production probably carried some semantic content, even if its primary goal was formal in nature.

4. Variety of Activities

The fourth study examined the effects of activity variety on language learning. To examine this question (8) lessons were taught in each of two schools, one providing third grade and one, ninth grade, subjects. In each pair of lessons, the first nine minutes were nearly identical, but during the last half of each lesson we employed either one type, of four types, of activities. As a means of identifying lesson content and number of semantic practice activities, we used the following identification system: for the first teaching point (TP_1), Lessons EA4 (four activities) and EAA (*Do and Say* Activity); for (TP_2), Lessons AA4 (four activities) and AAD (*Dramatization*); for (TP_3), Lessons FA4 (four activities) and FAP (*Picture Description*); for (TP_4), Lessons VA4 (four activities) and VAG (*Guessing Game* Activity). Test data obtained following each of these lessons served as the basis for comparing levels of learning. Most of the statistical variation was due to differences in the mean scores of third grade learners, since ninth grade groups had identical results. The effects of particular classroom behaviours on learning may be age-specific.

After scoring and tabulating the test scores, we did a number of analyses. Initially we computed t-tests using test scores from our sample in each school to determine levels of significance of differences in test date when students participated in lessons having either one activity or four activities during the semantic practice phase of the lessons. In School DE-37, the elementary school, we obtained a $p < .45$ for the listening comprehension scores and $p < .05$ for oral production scores, both favoring subjects involved in classes having only one type of activity in the final nine minutes. When the instructor taugh similar lessons to ninth grade learners, in School DH-23, we obtained $p < .40$ for aural comprehension scores and $p < .20$ for oral production scores, the former favoring those exposed to four semantic practice activities and the latter those involved in one such activity. With aural comprehension test scores we saw opposing tendencies, while with oral production scores we observed that students of both age groups generally performed better on tests when the teacher employed only one semantic practice activity.

To obtain more specific, and possibly more accurate information, we computed separate analyses of covariance on test scores received from pairs of lessons in both the elementary and secondary schools. In Table 38 we see the statistical results.

TABLE 38

Comparison of Levels of Significance of Effects of Variety on Learning

Teaching-Points Taught	LEVELS OF STATISTICAL SIGNIFICANCE			
	Aural Comprehension		Oral Production	
	Third Group	Ninth Group	Third Group	Ninth Group
TP_1	$p \leq .07$	$p < .05$ (4)	$p \leq .69$	$p < .97$
TP_2	$p \leq .05$ (4)	$p < .01$	$p < .43$	$p < .22$ (4)
TP_3	$p < .57$	$p \leq .16$ (4)	$p \leq .01$	$p < .01$
TP_4	$p \leq .48$ (4)	$p \leq .14$	$p \leq .37$	$p < .28$ (4)

(4) = Superior Learning With Four Activity Types

As Table 38 indicates, there was considerable variation in levels of statistical significance both within and between the different age groups.

4.1 *Third-Grade Learners*

As we stated earlier, statistical analysis (using t-tests) of data gathered after each of eight language lessons had been taught to third grade learners, resulted in $p < .45$ and $p < .05$ for aural comprehension and oral production respectively. In both cases, subjects involved in lessons having one type of semantic practice activity tended to perform better. To find possible explanations for these figures, we categorized and quantified classroom behaviours.

As means of interpreting variations in mean scores on tests of aural comprehension, we first considered behaviours whose essential goal is to get lesson content into the minds of learners, that is, presentation behaviours. When we totalled the duration and frequency of these behaviours in lessons having four semantic practice activities, we found that semantic presentation behaviours occurred 79 times and lasted a total of 292.4 seconds, and formal presentation 100 times and lasted a total of 152.7 seconds, for a grand total of 179 occurrences, lasting 445.1 seconds. In the four lessons having one semantic practice activity, the instructor employed semantic presentation behaviours 73 times, lasting 293.6 seconds, and formal presentation behaviours 105 times, lasting 159.4 seconds, for a grand total of 178 presentation behaviours, having a total duration of 453.0 seconds. From these figures, we see that the distribution of presentation behaviours was similar for the two types of lessons taught.

Subsequent categorization and quantification of more refined categories of presentation behaviours confirmed the similarity in distribution of behaviour classifications during the presentation phase of each lesson type.

Since only minor variations existed in the distribution of presentation behaviours, we quantified other teacher didactic behaviours. The analysis of semantic behaviour categories, especially the SA, SC and SS classifications, indicated that these behaviour types had 37 more occurrencies, lasting a total of 59.4 seconds longer in lessons having one type of semantic practice activity. Such a difference might possibly explain, to a large extent, how we obtained a $p < .45$ for aural comprehension test scores favoring students involved in these types of lessons.

Moving from didactic to language behaviours, we classified and quantified teacher language use categories. Behaviour quantifications revealed that every category denoting second-language usage occurred more often and had a longer total duration in lessons having one semantic practice activity type.

The only language use behaviour to have a longer duration and to occur more often in the lesson having four types of semantic practice activities was mother tongue (MT) usage. A great deal of this MT usage was employed as classroom administration because, on an average of once every 2.25 minutes during the last half of these lessons, the instructor was required to provide directions to learners on how to perform succeeding activities. Our analysis of articulation behaviours revealed that the CA sub-category occurred more often and lasted longer in every lesson having four kinds of semantic practice activities. The MT uses, then, often did not convey a great deal of information regarding the semantic content of the lessons.

While still considering Articulation behaviours, it is possible that the duration of OA behaviours, which consisted mainly of silence, may have indirectly influenced aural comprehension scores. During lessons containing four types of semantic practice activities, OA had a total duration of 1065.5 seconds, while in the other lessons it lasted a total of 955.1 seconds, for a difference of 110.4 seconds. During this time basically no teaching or learning activities took place. This loss of learning time might possibly have been reflected in learners' test scores.

It may be, as Krashen (1978) indicated, that learner utterances serve as input, and, as such, influence comprehension. When we quantified learner action and speech behaviours, we found that most categories of second-language speech had a higher distribution in lessons having one semantic practice activity type (see Table 39).

When we totalled the figures shown in Table 44 in the two different types of lessons, we found that students in the groups having four types of semantic practice activities used the target language 480 times, lasting 709.3 seconds while those involved in lessons having only one such type of activity employed it 523 times, lasting 803.6 seconds. When we condider this information, plus the quantifications of instructor behaviours, we see that the students who generally had higher aural comprehension test scores were normally those who had higher learner, plus total, second-language input.

The distribution of student second language speech might also have played a role in producing significantly different mean test scores ($p < .05$) on oral production tests, favoring subjects involved in groups having only one type of semantic practice activity. As we have already noted, students in these groups used the target language more frequently (43 more occurrences), and for a longer total duration (94.3 seconds longer). This additional distribution of second language use, plus higher teacher input, might have helped produce higher mean scores on oral production tests for subjects involved in lessons where the instructor employed only one type of semantic practice activity.

It is also possible that the duration and frequency of semantic-type student repetitions was an important factor affecting learning. Of course, one might argue that after presentation all repetitions were meaningful because the students already understood the content. Aside from that line of reasoning, action and speech categories DS, KS and LS, which in our lessons required actions, objects, or pictures along with speech, occurred 274 times having a total duration of 326.7 seconds in the lessons with four types of semantic practice activities. In the other four lessons they occurred 312 times and lasted a total of 414.7 seconds. If it happens that meaningful activities are most effective in producing learning, we would still have expected learners in groups having one type of semantic practice activity to have had superior test performances.

TABLE 39

Total Distribution of Action and Speech Behaviours in School DE-37

Behaviour Categories	ACTION AND SPEECH BEHAVIOURS			
	Lessons Having Four Semantic Practice Activity Types		Lessons Having One Semantic Practice Activity Type	
	Duration	Frequency	Duration	Frequency
DS	20.7	16	28.5	22
KS	200.7	160	274.5	190
LD	498.2	123	516.7	121
LR	105.3	98	111.7	100
LS	1.7	1	0.0	0
TR	12.4	7	8.4	4
RT	380.9	205	388.9	211
OL	102.0	72	92.1	77

The analysis of student language behaviour revealed that, except for phonetic errors, few learner miscues occurred in either lesson type. However, behaviour quantifications did indicate that students exposed to one type of semantic practice activity during the last half of the class had a total of 236 correct individual repetitions, for a total duration of 285.4 seconds; in the other lessons, learners produced 208 correct utterances, lasting a total of 243.6 seconds. Perhaps just as important is the fact that the same students had more correct independent utterances than the other group (57 compared to 40). Groups who used the target language more often, also produced more correct repetitions.

Research on human behaviour has often linked the terms *variety, motivation* and *boredom*. Stated simplisticly, investigators frequently claim that variety leads to increased motivation, but its absence induces boredom and resultant learning decrease. In the current study, because of the novelty of the subject-matter, the relatively short duration of the lessons, and the types of materials and activities employed, boredom was not a factor influencing test results. Even when the teacher employed only one semantic practice activity during the final half of a lesson, we found that, possibly because of the enjoyable nature of the activity learners had absolutely no difficulty concentrating on the activity for the time available. The monotony factor, then, did not enter into our interpretation.

With younger learners, the effect of varying the number of activity types, especially on oral production, seemed to be indirect, rather than direct. Third grade subjects tended to exhibit superior oral production skills, as measured by tests, when only one type of semantic practice activity was employed during the lesson. In this type of lesson, subjects normally had more opportunities to employ different categories of the target language, and tended to produce more correct individual utterances. This additional practice enabled subjects to better develop second-language speech habits.

Analysis of aural comprehension scores from third grade subjects seemed to indicate a relationship between input, especially learner input, and measured learning. In lessons where the distribution of instructor and learner utterances, especially those belonging to more semantic categories, was higher, subjects tended to score higher on listening comprehension tests. Again, the effects of varying activity types seem to be indirect rather than direct.

4.2 *Ninth-Grade Learners*

After analyzing test data and classroom behaviours in third grade subjects, we conducted identical types of analysis for lessons involving adolescent learners. Overall, t-tests yielded $p < .40$ and $p < .20$ for aural comprehension and oral production respectively. When the instructor employed four types of semantic practice activities, learners tended to perform better on aural comprehension; when he utilized one such activity type, they generally had slightly higher oral production scores. In order to explain these data, we analyzed language lessons.

In an attempt to find clues which might explain why the difference in mean level of aural comprehension occurred, we initially analyzed teacher didactic behaviours. Quantification of presentation behaviours revealed that the distribution of all categories, except PA, which was almost identical in both lessons, was marginally higher in lessons having four kinds of semantic practice activities. We found that in lessons having four kinds of semantic practice activities, presentation behaviours occurred 231 times, having a total duration of 365.9 seconds, while in the other lessons, similar behaviours occurred 219 times, lasting a total of 368.7 seconds.

Subsequent analysis of instructor repetition behaviours indicated that the teacher employed the target language 246 times for a total duration of 310.1 seconds in lessons having four types of semantic practice activities; in the other lessons, teacher repetitions occurred 268 times, lasting a total of 325.7 seconds. The more semantically-oriented categories (SA, SX and SS), however, occurred 187 times in the former lessons, and 188 times in the latter. With the type of distribution differences noted in various categories of teacher repetition, one would probably feel that the order of distribution variations might have been reflected in similarities noted in mean aural comprehension scores. Quantification of verification behaviours also failed to produce significant behavioural differences.

Since the distribution of instructor didactic behaviours was similar in the lessons being considered, we analyzed teacher language behaviours in order to ascertain whether or not types of language use could have been related to differences in learning. In total, except for the OU classification, various types of instructor language behaviours occurred 423 times and lasted a total of 565.8 seconds in lessons having four types of semantic practice activity; they occurred 439 times and lasted an aggregate of 556.2 seconds in lessons of the other type. While, overall, the distribution of these behaviours was similar in both types of lessons, the MT category occurred more often and for a much longer duration in lessons having four semantic practice activities. In that lesson type, these behaviours happened 227 times during a period of 629.0 seconds, while in the other, similar behaviours occurred 197 times and lasted 411.4 seconds. In both lessons, however, CA behaviours composed much of the MT category. In lessons having four types of activities during the last nine minutes, MT (CA) behaviours occurred for extended periods. Such prolonged periods of classroom administration decreased learning time, and might have diminished learners' level of aural comprehension.

Learner input as well as instructor input may influence listening comprehension. Lesson analysis indicated that, in lessons having four types of semantic practice activities, learning behaviours tended to have shorter durations. Analysis revealed that much of the variation in learning time could be accounted for by differences in the distribution of KS, LS and OL behaviours (see Table 40).

TABLE 40

Total Distribution of KS, LS and OL Behaviours Among Ninth Grade Learners

	SPECIFIC LEARNER SPEECH BEHAVIOURS			
Behaviour Categories	Lessons having Four Types of Semantic Practice Activities		Lessons having One Type of Semantic Practice Activity	
	Duration	Frequency	Duration	Frequency
KS	158.5	151	235.1	208
LS	45.8	38	0.0	0
OL	70.9	57	115.0	110

From Table 45, we observe that the two categories involving the target language had higher distributions in different lesson types. With the duration and frequency of KS behaviours being so notably higher in lessons having one semantic practice activity type, one might have anticipated superior aural comprehension scores from students involved in that kind of lesson. However, the presence of 110 OL behaviours (compared to 57 occurrences in the other lesson type), usually disruptive in nature, could have had a retarding effect on learning. Considering these facts, the overall influence of learner behaviours on aural comprehension levels may be viewed as questionable.

During our discussion of overall learner aural comprehension skill level, we have discovered that, while total instructor input was similar for both types of lessons (502 occurrences in lessons having one type of semantic practice activity and 515 in the other), learner target language utterances were more frequent and longer-lasting in lessons having one such type of activity (561 incidents compared to 527). However, the frequent occurrences of learner OL disruptive behaviours in that lesson type, might have had a negative effect on level of aural comprehension for students who participated in these lessons.

Oral production, as well as aural comprehension, may be affected by classroom behaviours. If we view language learning as habit formation, since students involved in lessons having one type of semantic practice activity tended to use the second language more often, we might have expected these students to generally perform better on oral production tests. Besides that fact, students in the same groups used the language more often with referential materials, 235 occurrences (for 269.3

seconds) compared to 218 (240.7 seconds) for the other students. Both total language usage, as well as language use with objects, actions and pictures, occurred more frequently in lessons having fewer types of activities.

When we examined individual student performance we found that students having one type of activity during a lesson produced more individual utterances containing phonetic errors, 128 compared to 89 (a factor not taken into account in our test results at this stage), but fewer of other types of errors (10 compared to 18). We found an equal number of correct responses (209 each) in each lesson type. Lessons having a greater variety of activity types contained more utterances classified as OK (152 compared to 137) and CC (20 versus 12), but the other lesson type had more correct independent utterances, 60 compared to 37. In every pair of lessons, SI was more frequent in lessons having one type of semantic practice activity. This latter factor may have influenced considerably the oral production scores.

In sum, analysis of test data from ninth-grade learners, indicated that, overall, adolescents' aural comprehension skills tended to be slightly higher after involvement in four kinds of semantic practice activities. Oral production skills, however, had a tendency to be superior after learner participation in lessons having only one such activity type. Behaviour quantifications provide possible explanations for these results.

On both types of tests, however, performance could also have been influenced by a learner factor. For aural comprehension, one group always had superior mean scores; for oral production, the other learners tended to have higher test scores. It is possible, then, that performance on a particular test type could have been related to the learners themselves. Thus, for ninth grade learners the results are somewhat inconclusive.

The trend in second language classroom research has been to evaluate learning in terms of particular theories and methodologies. However, practitioners who adhere to a specific theory or teaching method teach in a particular way. Certain behaviours or activities are emphasized while others are de-emphasized. It is this utilization of specific behaviour types, which are characteristic of the chosen method influencing learning. A teaching method, then, will influence language acquisition only if it facilitates utilization of particular appropriate classroom behaviours.

The studies discussed above are based on the principle that classroom behaviours have an effect on the teaching-learning process. No attempts have been made to evaluate particular philosophies, theories or methodologies. Rather short-term experimental studies were conducted with a view to determining the effects of particular classroom practices on student learning. The polychronometric technique was employed to analyze the classroom interaction process of as many as 100 teacher and student behaviours. This type of analysis enabled us to establish relationships between specific behaviour distributions and different levels of learning.

This, coupled with the experience described in the foregoing chapters, illustrates the possibility of measuring the value and efficiency of classroom language teaching.

Conclusion

MEASURING LANGUAGE TEACHING EFFICIENCY

1. **Syllabus Analysis, Method Analysis and Lesson Analysis**
2. **The Development of Lesson Analysis**
3. **Polychronometric Techniques in Lesson Analysis**
4. **Cost-Benefit Analysis of Language Lessons**
5. **Toward More Efficient Language Teaching**

In measuring the efficiency of foreign language teaching, there are several layers in the hierarchy, between the education authority and the individual learner, that have to be taken into account. There is the curriculum, the syllabus, the methods and materials, the teachers and their skills, and the individual lessons. In the last analysis. it is the nature and amount of the language acquired by the individual learner as seen and heard in his or her on-going performance that is the final measure of the efficiency of language teaching. It is that measurement, as placed in an hierarchical context of classroom teaching, which will be addressed in this conclusion to our book.

1. Syllabus Analysis, Method Analysis and Lesson Analysis

Most of the language teaching which takes place in schools today is done according to a local, regional or national syllabus which defines what should be taught and learned at each level of the educational system in which a foreign language is taught. The content of this syllabus is the first possible cause of failure in foreign language learning. The objectives may be unclear as statements of which skills are required at each level and the extent of each requirement. What is required may be much more than can reasonably be achieved in the time allotted. Specifications may be too vague to be implemented. Or the syllabus may require the learning of skills that are little used and soon forgotten.

Teachers, however, seldom teach directly from a syllabus. They generally teach from a teaching manual or textbook which has been approved as being in conformity with the requirements of the syllabus. All or part of the textbook and teaching materials may also be in the hands of the learners. It is this, therefore, and not only the syllabus that may be at fault. The textbook might not reflect the syllabus. Its content may be unsuitable for the type of learner to which the teacher must teach. Its method of presentation may be unsuitable to the age characteristics of the learners or to their interests, abilities or background. The textbook may also be unsuitable to the type of teacher who has to use it. Its use may require linguistic or professional skills which the teacher does not possess. Or it may be inappropriate to the conditions under which it has to be used large classes, few contact hours, low priorities and the like.

The fact that a method, its texts and materials, and the syllabus on which it is based are indeed suitable does not imply that it is effective as a language teaching method. It may be deficient in the choice of what it includes, the order in which it teaches the items included, in the way these are presented, or in the type and amount of drills and exercises (Mackey 1965).

In the analysis of this input, however, a distinction must be maintained between the materials of the method (textbooks, workbooks, tapes and films) to which the learner is exposed and the teaching activity to which the class is subjected. Variables in methods have been isolated and many have been the object of quantified and even automated analysis. On the basis of such analysis, detailed studies of method have been made on each of the four didactic components in answer to the four basic questions: what, when, how and how much. Comparative studies have been made of the effectiveness of selections used in different methods. Theoretical and experimental studies have also been completed on the effects or different types of gradation in formal language learning. Such studies in method analysis can be undertaken even before the materials get into the hands of the teachers. It is only after such analysis has been completed, however, that one is able to determine which deficiencies in the language learning process are attributable to the method and which depend on the way it is used. In sum, much can go wrong with language teaching in the educational system even before it gets near the classroom.

No matter how well designed it may be, once a method is in the hands of the teachers there is no guarantee that it will be used as intended. It is the use of the materials and not the materials themselves that may be the reason for the success or failure of the language teaching process. In analysing this, the first and most obvious thing to find out is how the teacher intends to use the method, in the way of omissions, additions and modifications. What use will the teacher make of the formal and semantic possibilities of the materials (Mackey 1965). In the analysis of language teaching activity as input, therefore, the first thing is to determine the extent to which the learner is exposed to the materials, either directly (reading and listening to recordings) or through the intermediary of the teacher - how much, how often and for how long. In most current approaches to language teaching whether they be direct, audio-lingual or audio-visual, most of this exposure takes place in class and in the language laboratory, if the objective includes the acquisition of the spoken language. It is therefore observable and capable of being studied as behaviour.

Since most of this behaviour appears in the form of language lessons, it is the language lesson which constitutes the appropriate unit of analysis. In analysing a lesson a basic distinction must first be made between what the teacher intends to do and what the teacher actually does. That is, there is a difference between a plan and a performance.

Much can be learned by starting with an analysis of a teacher's lesson plan. The first thing to look for is what the teacher intends to achieve under known conditions within the given period of time. What are the teaching points and how are they related among themselves and made to tie in with what has gone before and what will come after? What procedures does the teacher intend to use to present new forms, to make these clear and to convert them into habits? What is the proportion and order of the techniques to be used?

A plan, however, is not a lesson, and in the last analysis, it is the performance of the teacher which will determine the quality and the quantity of the learning. In analysing this performance lesson by lesson it is important to distinguish between what is taught and how it is taught. If the plan can deviate from the method, the performance can deviate from the plan. In this performance we can find omissions, additions and modifications of what the teacher had intended to say and do. In analysing how things are taught it is also important to relate the context to the teaching conditions. Is the layout of the class, for example, appropriate to the techniques to be used? Are the techniques appropriate to the type of learners - their age, ability, motivation and background? How are the presentation techniques related to the techniques of repetition? These are the sort of questions relevant to the analysis of the content of the lesson (Mackey 1965).

After having analysed a number of lessons in this way, one notices a difference in results, which can only be attributed to differences in the teaching and learning behaviour which takes place in the classroom. Not only are the components of such behaviour different, but what is more, each component differs in both duration and frequency of occurence. One ends up therefore with the basic problem of determining how much of which type of behaviour is responsible for what difference in language learning. In sum, one is faced with the necessity of quantifying both language teaching behaviour and language learning performance in order to find out why the learners are not acquiring what they are supposed to learn.

In trying to find out why something is not working as it ia supposed to - from troubleshooting a stalled car to debugging a computer programme - the important thing is to look first for the obvious. It would be foolish to dismantle a stalled car only to find that the battery has simply run down. It would likewise be a great waste of time to go directly into quantitative lesson analysis before all the qualitative variables had been fully evaluated. All the more since these, including as they do the professional and linguistic qualifications of the teacher, are often directly observable and readily evaluated. It would seem wasteful, for example, to find out after painstaking quantification of classroom behaviour why the learners in a given class never succeed in learning to speak the foreign language properly when a few minutes of direct observation makes it obvious that the teacher speaks no better. Analysis for the sake of analysis is not only wasteful; it can lead to some impossible aberrations from everyday common sense.

Once analysis is justified, however, it must be undertaken in such a way as to yield the information expected, according to a known level of reliability. This means that the analysis must have purpose, direction, method, procedure and techniques independent of person, place and time, if observations and quantification of classroom behaviour are to be valid.

2. Development of Lesson Analysis

Within the domain of educational research, the observation and quantification of classroom behaviour is nothing new. It can be traced to the beginning of the century, at least as far back as the year 1914, when Horn experimented his technique for gauging the degree of pupil participation in a lesson. This technique consisted in marking a seating chart with circles and squares representing actions and responses (Medley & Mitzel 1967). By 1928 more categories and more symbols had been added by those using the technique (Puckett 1928). The following year, Barr was able to report on the characteristics of good and poor teachers of social studies (Medley & Mitzel 1967). The same year, classroom observation was used by Thomas to study the relationships between pupils in a nursery school (Medley & Mitzel 1967). In 1945, such kindergarten observation was used by Anderson and Brewer to analyse the classroom personalities of teachers (Anderson & Brewer 1975). And three years later it was used by Whithall in a study of the social and emotional climate of a classroom.

In the sixties, a workable system was proposed for the measurement of the use of the different language skills in the classroom. This system of observation, proposed by Jarvis, contributed to the objective analysis of language teaching behaviour by making a basic distinction between real speech and drill speech, the real speech taking place in either the native or in the foreign language (Jarvis 1968). Each was recorded in time-units of ten to fifteen seconds.

Since the sixties, and even before, the analysis of classroom behaviour seems to have concentrated on the quantification of interaction between teacher and learner. In 1956 Cogan suggested a new design for the study of teacher-learner interaction (Cogan 1956). A decade later, Amidon, in collaboration with Hunter, Flanders, Hough and others came out with a series of books and articles devoted to interaction analysis as a way of improving the effectiveness of teaching (Amidon & Flanders 1967). There were applications to the teaching of specific subjects such as reading, geometry, physics and other sciences.

One of the most elaborate systems of general interaction analysis of teaching behaviour was developed in the sixties by Flanders for the purpose of quantifying the qualitative aspects of verbal behaviour in the classroom. The system was based on the assumption that the teacher influences the class through spoken statements, any non-verbal behaviour, simply being a reflection of the verbal. Statements could among other things be used to provoke student questions, responses and verbal interaction in general. The ten categories comprising the Flanders system are divided into those representing teacher talk and those related to student talk. Teacher talk includes accepting feelings, praise and encouragement, accepting or using ideas of students, asking questions, lecturing or orientation, giving directions, and criticising or justifying authority. Student talk includes limited talk, unlimited or initiated talk, confusion, or silence.

Teacher categories were further subdivided into those which indicated indirect influence upon the learner or the tendency on the part of the teacher to maximise the freedom of the students to respond (Categories 1-4), and those which indicate direct influence or the tendency on the part of the teacher to limit the learner's freedom to participate (Categories 5-7). In a later work, Flanders renamed these two dimensions "initiation" and "response", and interaction analysis became the means of establishing the balance withing the classroom between these categories (Flanders 1967, 1970).

The tallying of responses according to these ten categories is done by trained observers who note the corresponding category number (1-10) every three seconds - after deciding which of the ten dominates the three-second period. Numbers are written in sequence in a column, and each, except the first is subsequently paired with the number immediately preceding it and the one just below. Each pair is then entered, as a tally, in one of the hundred cells of a ten-by-ten matrix. The results are expressed in terms of percentage of teacher talk and student talk, with indirect/direct ratios representing the type of teacher influence in the classroom, or alternately, the degree of freedom afforded the students during the lesson.

This system was in turn adapted for the analysis of the teaching of particular school subjects. In the field of foreign language teaching in the United States it was first adapted by Moskowitz who maintained all the categories and supplied guidelines to foreign language teachers for encoding and classifying the activities in the language class (Moskowitz 1968).

As the system was used in the training and re-training of foreign language teachers, it underwent further modification. Need for another language dimension in the language class, led Wragg to differentiate between the use of the foreign language and the use of the native language (Wragg 1970). To do this he made use of the first ten of Flanders categories for one language and an extra ten for the same categories in the other language, thus producing a twenty-by-twenty matrix comprising four cells on the bases of combinations in the classroom use of the foreign language (FL) and the native language (NL): (NL & NL), (FL & NL), (NL & FL), (FL & FL). In 1970, Moskowitz carried her adaptation of the Flanders system a step further by including the kinds and amounts of student and teacher talk in a system which became known as "Flint" (for Foreign Language System of Interaction Analysis) when in 1966 it first appeared as an adaptation of the Flanders system of interaction analysis in the teaching of school subjects (Moskowitz 1971).

In addition to the above-mentioned categories, the FLint system is designed to take the following into consideration: a) the tendency of the teacher to deal with errors by correcting students with or without rejection (Categories 7a and 5a), b) whether the teacher criticized the student's behaviour, or the responses (Categories 7 and 7a), c) the amount of choral response (Category 8) whether confusion is work oriented or not (Categories 11 and 11a), e) how much levity there was in the classroom (Categories 2a and 12). To the fifteen categories can be added a further category ('e') to indicate that the behaviour took place in English, in this case, the native language of the learners. For example, if the teacher asks a question in the learner's native language a "4e" is recorded; if the entire class responds in the native language, an "8ae" is noted. To any of these categories an "n" may be added to indicate non-verbal behaviour. For example, if the teacher confirms that the student's response was correct, this is noted as "2n" (Louisy 1975).

A further refinement of the adaptation of the Flanders system was made for specific levels of language teaching, notably in the elementary school. This was the contribution of Rotfarb who simply subdivided teacher talk into native and foreign and student talk into individual and group thus producing a system of 13 categories: 1. modeling, 2. giving directions, 3. asking direct questions, 4. guiding structure drills, 5. rephrasing pupil responses, 6. reacting to pupil performance, 7. lecturing, 8. reading, writing and spelling. Student talk: 9. responding, 10. answering direct questions, 11. student-initiated talk, 12. reading, writing and spelling, 13. silence or confusion (Rotfarb 1970).

The purpose of the system was to reveal the nature and extent of teacher talk, student talk, individual and group response and of the use of each language, in this case English and Spanish. Each response was therefore coded by category represented by its number on the above list and by a letter (T for teacher talk and I or G for individual or group response). If the interaction occurred in the foreign language (Spanish) an "Sp" was written. For example, teacher talk in Category 2 giving directions in Spanish would be coded 2TSp. The system was first tested in California and later applied in Miami, even in some of the bilingual classes which at that time were rapidly growing in number and importance.

Most of these different systems of interaction analysis have been used for the purpose of improving the quality of classroom teaching by inducing the teacher to increase the amount of interaction with the learners. Much emphasis is placed on the social and emotive climate of the class, including the improvement of the teacher's teaching manners.

The system of interaction analysis, however, has not been without its critics. Some have observed that this has developed into a sort of new educational religion whose chief objective is to provide the teachers with skills and techniques of good classroom manners so that they may behave more decently toward students. Although this criticism would seem both unfounded and unfair, the conclusion of investigators who have used interaction analysis extensively in language teaching research would seem more to the point. Aside from its value as a teacher-training device, it has been noted by researchers that while half the categories may be unused, what remains is not sufficient to permit the coding of what may be relevant to the language-learning process. More categories were particularly needed to distinguish the various types of student verbal and non-verbal behaviour. The discovery of these deficiencies, however, is not surprising since the system was not primarily designed as an instrument of research into foreign language learning. It was designed for the analysis of interaction in the teaching of school subjects, that is for a type of behaviour where teacher and student talk *about* the object of the teaching, which is primarily not a foreign language. When it is language as a subject, as in the teaching of the mother tongue, it is really language about language, that is, linguistics in the form of grammar rules or word-study. It is not language as a new skill in which talk for the sake of talk may be the dominant element. Interaction analysis as an evaluative technique is based on the assumption that most of the influence of a teacher is expressed through verbal behaviour. This type of behaviour (language) is the means to the end (learning). In foreign language classes, however, the means to the end (language) may not always be limited to the use of language, but may also include the use of pictures, objects, actions and situations. The analysis of the language teaching and language learning process cannot be entirely objective if made within a preconceived theory of learning based on the study of the acquisition of knowledge as opposed to the mastery of a skill.

It is for these reasons that some research into language teaching and language learning behaviour took place, not in the context of a theory of learning such as assumed in the systems of interaction analysis, but on the basis of direct observation of the process of foreign language learning and teaching. It was rather in the context of the analysis of language teaching the purpose of which was to provide techniques for the study of the relation between classroom behaviour and foreign language learning. It was this need which led to the study of time variables in classroom behaviour and to the development of polychronometric techniques of lesson analysis.

3. Polychronometric Techniques in Lesson Analysis

In order to study the relationship between classroom teaching and language learning, it was necessary to develop techniques of analysis and quantification of classroom activities that could have some effect on the learning of a foreign language. This implied the identification of categories of specialized didactic (teaching and learning) behaviour on the one hand, and types of language behaviour on the other. Since the effect of each type of behaviour could depend on the amount used, this implied not only the objective identification of each type of event in a classroom lesson, but also an accurate quantification of how long and how often it took place. This turned out to be no simple matter.

As we have seen, the number and speed of events that took place in certain types of language lesson required the development of some sort of adequate instrumentation. Since only a limited number of categories of behaviour could be noted at any one time by an observer, there was the likelihood that some sort of recording of the lesson might be necessary in more thorough investigations, which went beyond the most general categories. Finally there was the requirement to develop valid methods of computing and interpreting what had been observed. In sum, the problems were those of instrumentation, categorization, computation and interpretation (See Chapters 1 & 2).

Once these problems were solved and techniques developed for demonstrating the relations between teaching and learning, it was possible to think of a language lesson as a goal-oriented set of behaviours the efficiency of which could be measured as a function of the stated objectives. The outcome could then be evaluated as benefit in terms of learning achievement related to cost in terms of time and effort expended, in other words as the result of a cost-benefit analysis.

4. Cost-Benefit Analysis of Language Lessons

In this type of quantification, the basic question is: how much language learning does the teaching achieve in relation to the investment made. Investment, or cost, can be class time, class time per teacher, or class time in relation to the teacher-student ratio. Benefit can be defined as the actual amount of language learning as distinguished from discourse learning or memory, that is, the generation of acceptable sentences from a system, rather than the mere repetition of sentences already presented. If the objective is oral skill, the benefit is assessed as a demonstrable use of this skill in generating sentences correctly and independently.

Since lessons vary in length, purpose and design, cost-benefit analysis of them can be valid for the language teaching process only when made on a representative sample of such lessons. Each lesson is then analysed as a specimen of the sample and the total results are averaged.

If the class size is over ten and if polychronometric techniques are to be used, an audio or audio-visual record will be required. If the record is audio, some notation will be required to identify one speaker from the other. This can be done during the recording process by marking sequence numbers on a large-enough seating chart. If videotape recording is used one camera should focus on each student speaker and remain until the utterance is completed. In using polychronometric techniques to determine the learning benefit of a lesson expressed as production of what is being taught, one of the keys is therefore allotted to each member of the class. If there are more learners than there are keys, the analysis can be made from the recording in sets of ten learners, or fewer. When a learner produces an acceptable sentence of the type required by the analysis, the observerd touches the key which has been allotted to the speaker. When another learner does likewise another key is depressed, and so on, until the end of the lesson (See Chapter 1).

At the end of the lesson there will be figures on both the timers and the counters. Those on the timers represent elapsed time the total of which should be equivalent to the length of the lesson. Those on each of the counters indicate the number of complete and acceptable utterances generated by the learners to whom the corresponding key had been allotted. The cost-benefit is calculated first from both sequential analysis and from the sort of lesson profiles illustrated in the foregoing chapters.

4.1 *Sequential Analysis*

Sequential analysis will reveal the points at which each category is use, for how long it is used and how often it recurs. By way of example, let us consider the sequential analysis of an elementary

English lesson of twenty minutes duration given to a class of ten learners, all of them between the ages of nine and ten. The lesson has been recorded on videotape and analysed in sets of ten categories in the hierarchical order indicated above. Examples of the results for the first four runs appear in Figure 69. The time-line in this figure appears in minutes for the general analysis and in seconds for the break-down of the immediate constituent categories of articulation, presentation and repetition. The minute line is further subdivided into units of twelve seconds each. The sigla on the left (AR, SP, etc.) are identical to those listed in the above chapters. Two layers of sequential analysis are shown: 1. Didactic Behaviour (Teaching and Learning), and 1.1 Teaching (Articulation, Presentation and Repetition). Segments of the continuous line (---) opposite the sigla representing the categories indicate the length of time used by each category at the points in the lesson corresponding to that part of the minute or second line immediately above. The broken line (---) represents the time spent in the use of the first or native language of the class. Also opposite each category, but at the extreme right of the time line, are two columns of figures. The first represents the number of times the corresponding category appeared in the lesson (frequency of occurrence); the second represents the amount of time in seconds during which that category of behaviour appeared. For example, the number 14 opposite AR on the far right in the FRQ (frequency) column indicates that articulation activities such as linking, checking pronunciation and the like occurred 14 times; the figure to the right of this, 276 in the next column DUR (duration) indicates that these activities took up 276 seconds of the lesson. When further analysed into its components, however, this AR category is shown to have comprised a number of different activities each of which is further analysed as a subcategory in a later analysis. Each of these was analysed as a separate category during second runs of the lesson. Such second-run analysis is plotted here, by way of example, for three sets of categories: AR (articulation), P (presentation) and R (repetition). Here the sequential scale represents real time in seconds. One component of AR, for example, semantic verification (coded SV in the left hand column of sigla) occurred five (5) times in a total of forty-two (42) seconds, as indicated in the column on the right (See Figure 69).

On the basis of the representation of these time segments, it is possible to obtain a picture of the lesson as it took place, the picture being made up of the occurrence and recurrence of different categories of behaviour. For comparative purposes, however, it is a different picture that is required, one that integrates these data into quantitative profiles of time use and distribution.

4.2 *Integration Into Lesson Profiles*

Lesson profiles are obtained by integrating the data. By adding together all the line segments opposite each category in Figure 69, we get another line which represents the total amount of time (T) devoted to that category of behaviour during the whole lesson. It therefore represents a certain proportion of the lesson. For example, the fourteen segments of the AR line in Figure 69 occupy a total of 276 seconds, including silences, that is more than a four-minute portion of the total time of 20 minutes represented by the length of the horizontal time line in the profile on Figure 70. The second segment of the duration line is occupied by a representation of the time spent by the teacher on semantic presentation (SP) corresponding to the total time in seconds 123 indicated in the last right-hand column in Figure 69; and so on, for the segments FP (75 seconds), FR (30 seconds) and SR (132 seconds), followed by those representing time taken up by the learners: AC (51 seconds), SH (255 seconds), SA (194 seconds), RE (nil), and CH (64 seconds), until the entire total of time (T = 1200 seconds) of the 20-minute lesson has been accounted for on the horizontal axis in terms of teaching and learning behaviour.

As for the vertical axis, it is used to indicate language distribution within each behavioral segment, marked as a proportion of percentage of the whole. For example, in the activities grouped under articulation (AR), the mother-tongue was used mostly in classroom administration but occasionally to get drills under way. The mother-tongue (L_1) occupied 88 seconds of the 276, or 33%, the second language (L_2) 40%, and the rest, 27%, was silence. These proportions are indicated in the horizontal cuts in the vertical axis. The space between the two axis is reserved for number of occurrences of each type of behaviour occupying that segment of time indicated on the horizontal axis.

Each occurrence is represented by a line within the segment - a diagonal line for the first language and a horizontal one for the second language. For example, the mother tongue was resorted to six (6) times during the activities grouped together as articulation (AR); this is represented by the six diagonal lines in the AR segment. For the same activities, the second language was used in eight (8) segments; this is represented by the eight perpendicular lines in that portion of the AR segment representing the time during which the second language was used for these activities (See Figure 32).

In the same way the more detailed second-run profiles are constructed. Three examples of these (articulation, presentation and repetition) are given in Figures 31 and 32. Since each adds up to a shorter time line, it is possible within the space provided here to mark the horizontal axis in seconds of real time. Similar detailed profiles can be made for third and fourth-run sets of behaviour categories to represent such variables as feedback techniques and the several procedures used to elicit speech.

There may be need, however, to base such lesson profiles on the results of sequential analysis the automatic recording of which requires more elaborate instrumentation. One has simply to take the frequency and duration figures from the counters and timers and plot them directly on the lesson profile matrix. Here, for example, are the frequency and duration totals for other sets of categories not included in Figures 31 and 32.

In sum, lesson profiles permit us to see and compare at a glance the duration and frequency distribution of types of behaviour such as those computed above. These profiles do not in themselves constitute evaluations of the lessons. Such evaluation that can be made will depend on the interpretation of the presence and proportion of each of the categories of behaviour which make up the lesson.

FIGURE 31

Quantified Behavioral Sequences of an English Lesson

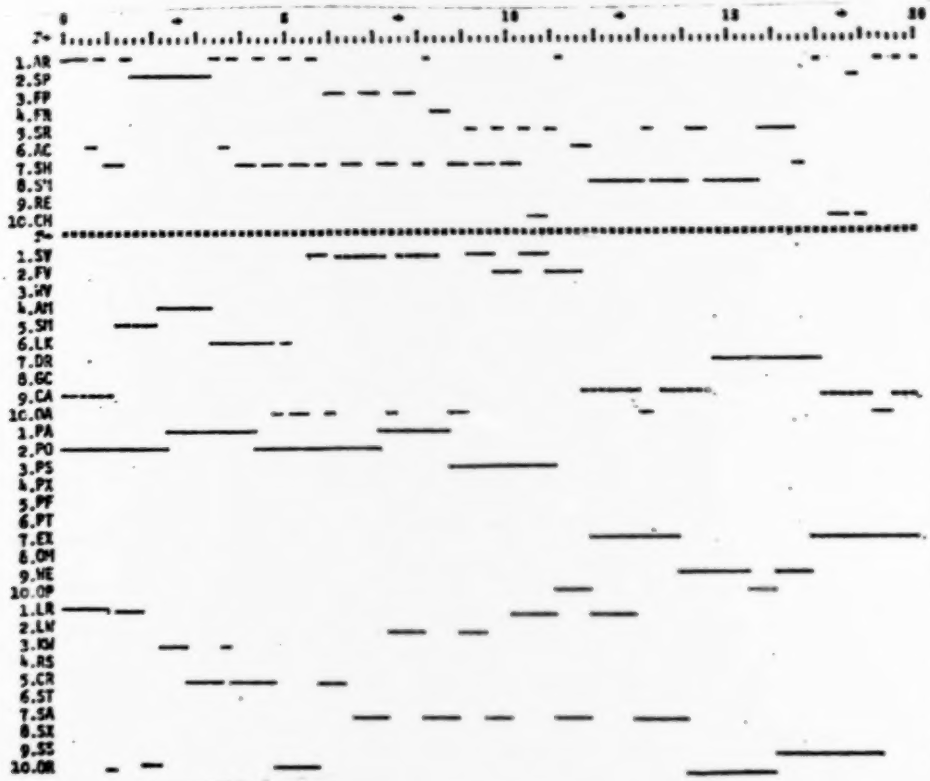
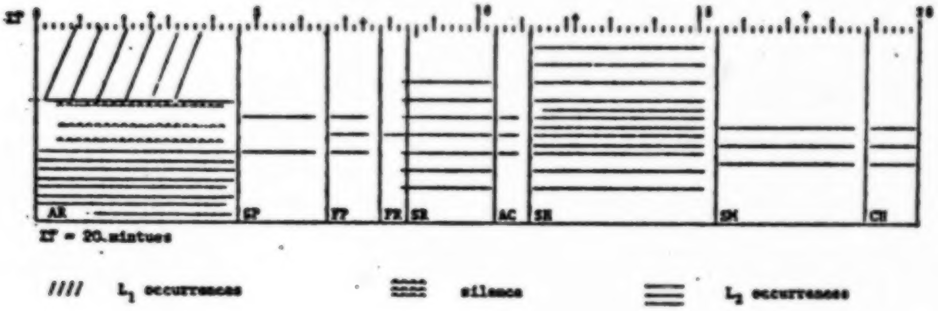


FIGURE 32

Behaviour Distribution Profile of the Lesson



1.1.1.1 Verification

	FRQ	DUR
1. PT	1	5
2. DO	2	13
3. AN	4	17
4. VR	0	0
5. FR	2	5
6. EM	1	3
7. RE	0	0
8. VB	0	0
9. CO	2	6
10. OW	1	2

1.1.2.1 Actions

	FRQ	DUR
1. VB	0	0
2. PB	0	0
3. PX	2	3
4. PO	0	0
5. MO	3	12
6. MV	2	6
7. MI	2	7
8. IA	4	5
9. JOH	0	0
10. OC	3	2

1.1.3 Speech through Speech

	FRQ	DUR
1. PP	6	6
2. QA	5	3
3. CC	7	12
4. OC	0	0
5. OU	0	0
6. RD	0	0
7. DE	0	0
8. DM	0	0
9. TT	0	0
10. [illegible]	1	3

1.2.1 Action and Speech

	FRQ	DUR
1. DS	19	193
2. RS	20	75
3. LD	6	24
4. LR	18	69
5. LS	12	71
6. LW	0	0
7. TR	0	0
8. RC	0	0
9. RT	21	83
10. OL	11	86

2.1 Language Use

	Utterances
1. DE	28
2. ER	16
3. MT	19
4. GD	0
5. QL	39
6. CS	9
7. IT	18
8. CP	30
9. CG	21
10. OU	26

2.2 Language Performance

	Utterances
1. PE	13
2. GE	11
3. VE	7
4. PG	21
5. PV	8
6. GV	13
7. AL	19
8. OR	11
9. CC	9
10. SI*	12*

4.3 *Interpretation*

Interpretations made possible through the above types of quantitative analysis make it possible to come to certain conclusions about the quality of the lesson. The analysis and the calculations permitted by the data can also tell us something about the value of the lesson as a goal-oriented activity. These qualitative characteristics of the lesson can be computed through the use of indicators of the qualities which one may seek or hope to find in a lesson. One may wish to find out, for example, whether one lesson or a sample of them compared with other lessons was more or less lively, interesting, intense or effective. To assess such qualities one needs to make use of reliable quantitative indicators of activity, variety, intensity and performance.

One measure of liveliness as we have seen is activity as measured by the number of changes that take place in types of behaviours over a given petiod of time. This activity can be expressed as the ratio of time (T) in minutes over frequency (f) of occurrences (tokens), so that,

In the above case, using only the first-run totals as indicated in Figure 69: $A = \frac{\Sigma f}{T}$

$$\Sigma f = 14 + 2 + 3 + 1 + 7 + 3 + 12 + 3 + 0 = 48$$
$$T = 20 \quad \frac{\Sigma f}{T} = \frac{48}{20} = 2.1$$

This can be interpreted to mean that behaviour in this lesson changes in type at least twice a minute, on the average. To obtain the maximum rate, data from second and third-run analysis would have to be used.

We have also seen that interest-promoting activities are characterized by their variety. This variable can be measured by counting the number of different didactic activities (types) having taken place during the lesson. These can be expressed as a ratio of the total (Σ) number of categories (c) of behaviour over time (T) in minutes. So that, $V = \frac{\Sigma c}{T}$. In this case, using only the data from the basic, first-run analysis, where $\Sigma c = 9$ and $T = 20$, $\frac{\Sigma c}{T} = \frac{9}{20} = .45$. Again the maximum rate would come from the data based on second and third-run analysis.

One indication of the intensity of the learning activity is the number of things the learners have actually said and done during a given period of time. In a language lesson this can be expressed as total numbers of didactic actions (ΣA) and utterances (ΣU) of the learners, both as individuals (IU) and as a group (GU) over the period of time (T) taken by the lesson. That is, $I = \frac{\Sigma A + \Sigma U}{T}$

Here, $IU = 126, GU = 21$, and $\Sigma U = 147$
$\Sigma A = 25$ and $T = 20$. So that $I = \frac{\Sigma A + \Sigma U}{T} = \frac{25 + 147}{20} = 8.6$

Not all actions and utterances included in intensity indicators, of course, will be complete or correct. Nor does the intensity of the learning guarantee the effectiveness of the lesson. Effectiveness has to be assessed in terms of performance.

Language learning performance can be measured in terms of correct utterances, proportion of errors, in type of utterance and types of errors in type of utterance. One indicator is the total number of correct individual utterances (ΣU_c) performed during the time (T) allotted: $\frac{\Sigma U_c}{T}$. Another indicator is the proportion of correct utterances compared to the total: $\frac{\Sigma U_c}{\Sigma U}$. The type of error can also be used as an indicator of the type of performance by establishing the ratio between the total number of utterances (ΣU) and the total number of errors for each type: phonetic (ΣE_p), grammatical (ΣE_g), or vocabulary (ΣE_v), giving indicators $\frac{\Sigma E_p}{\Sigma U}$, $\frac{\Sigma E_g}{\Sigma U}$, and $\frac{\Sigma E_v}{\Sigma U}$

Each of these may be judged in relation to the task performed. The lack of error in a dependent utterance (U_d) such as found in a sentence that has been repeated or only modified by the learner may be judged as being less of a performance than is the lack of error in an independent utterance (U_i), that is, one that has been completely generated by the learner. So that the indicator $\frac{\Sigma U_i}{\Sigma U_d}$ is given greater weight than the indicator $\frac{\Sigma U_c}{\Sigma U}$. Weighting may also be attributed to the type of error. A pronunciation mistake in an independent utterance may not have the same weight as would have a lexical or grammatical error. So that we may need to distinguish individually, for each learner: $\frac{\Sigma E_p}{\Sigma U_i}$, $\frac{\Sigma E_g}{\Sigma U_i}$, and $\frac{\Sigma E_v}{\Sigma U_i}$

Not all the correct and independent utterances, however, will be based on what was taught in the lesson. Yet it is only on the basis of such utterances that we can judge how immediately effective the lesson has been, that is, how much learning benefit, as it were, the learners have derived from the amount of teaching invested, that is, provided the units of benefit are properly counted.

The delimitation and definition of constituent units is a problem facing all such types of computation. The basic question is: What is counted as what? This question faces every practitioner of applied statistics. Unless this question is settled at the outset within an acceptable range of error-which must also be measured - no comparability is possible between one set of observations and another. Even in the field of language analysis everybody does not agree on the basic units. In the practice of lexicometrics, for example, the question of what constitutes a word has been debated ever since man began counting them. The language unit to be delimited here is not the word but the utterance. Not any sort of sentence, but only the type which has been independently generated by a language learner on the basis of material learned during a particular lesson. In counting such sentences so that they will represent comparable units, one must first establish and state the criteria on the basis of which each and every one of the acceptable sentences can be distinguished from all and any of the others. Only as a result of the consistent application of such criteria can one go about making comparisons between results observed in different specimens of language teaching behaviour.

Here, therefore, are a few suggestions for establishing criteria for excluding or including sentences as acceptable results of language teaching and learning behaviour (S1): 1. The utterance must be adequate within its context. (Some contexts, for example, may call for only a two or three word response). 2. The sentence must be grammatically and lexically correct. 3. The sentence must be phonetically correct for the level at which it is used. 4. It must be independently generated. (This excludes imitative or repetitive use of sentences uttered by the teacher, repeated after correction, memorized as parts of prepared skits or set dialogues or as language laboratory drills). This is of paramount importance if what we believe to be computing is the amount of language learning as opposed to the memorization of discourse. 5. The sentence must include one or more of the items taught (new structures, words or syntax). This excludes sentences made up exclusively of items which the learner has previously mastered or is already supposed to have mastered - although these may indeed be counted when testing achievement as such. At this point, however, it is not necessary to take into account the number of new items used in the same sentence, that is, its density. For although high density performance may be a positive measure of the ability or intelligence of the individual learner, it may represent, in the method used, a negative measure associated with a lack of didactic skill. Other criteria, such as fluency, tempo and significance, may be added to satisfy a particular purpose, according to the rigour required at different levels of language learning. The important thing is that the criteria chosen be unambiguous, clearly stated and consistently applied. Indeed, if the results are to be used as a basis for changes in teaching practice or policy, it is important that the same analysis be repeated by the observer and also replicated by a sufficient number of persons to establish its degree of consistency.

Applying criteria such as these, acceptable sentences are checked-off opposite the name, initial or number of each of the participating learners. For example, in the lesson analysed above, six out of the ten learners produced twelve independently generated and acceptable sentences (S1) on the basis of the new teaching points taught in the lesson which was of twenty minutes duration ($T = 20$). These were tallied in the order of appearance as in the accompanying table (See Table 41).

The reason why the utterances are arranged ordinally by learner, is that greater weight can thus be given to class learning as opposed to individual performance. The virtuosity of one or two brilliant learners can later be evaluated as individual performances, some of which may in fact be attributable to extraneous circumstances such as language background, bilinguality and residence abroad which can be measured separately and as such. Since what is being measured is the learning of the class as a group, it is important to see how this learning is distributed. If the twelve acceptable sentences referred to above were uttered by the same learner, the class as a whole would provide less evidence of learning than if each of the learners generated at least one of the sentences. It is important that this be taken into account in the formula used to measure how much learning has taken place. The factors of the formula must include the number of learners having independently generated through the use of the newly learned point at least one acceptable sentence (S), the number of learners having generated two such sentences, the number of learners generating three, four, and so on. The formula must also provide for a relation between these numbers and the total number of learners in the class (N) and finally for the time (T) used in learning to generate such utterance.

TABLE 41

Distribution of Independently Generated Sentences in a Lesson

Learner No.	1st $x = 1$	2nd $x = 2$	3rd $x = 3$	4th $x = 4$	5th +...n(S) Sentence $x = 5..n$	
1	S					
2						
3	S	S	S	S		
4						
5	S					
6	S	S	S			
7	S					
8						
9						
10	S	S	–	–	–	
$N = 10$	$S_1 = 6$	$S_2 = 3$	$S_3 = 2$	$S_4 = 1$	$S_5...n = 0$	$\Sigma S_x = 12$

The total production of the class of acceptable and independently generated sentence, therefore, can be expressed as the sum (Σ) of S_x ($s_1 + s_2 + s_3 + s_4 + ... s_x$) averaged for the number (N) of learners in the class, that is, $\sum \frac{S_x}{N}$.

Since it is the learning of the class as a whole, rather than the performance of a few productive individuals that we wish to measure, the results must be weighted in such a way as to take into account the former while placin the latter in its proper perspective. This can be done by reducing exponentially ($x^2 - 1$) the value of each successive utterance made by the same speaker thus neutralizing any possible monopolization by one or two individuals. The addition of this parameter to the measurement of cost-benefit (CB) gives us the formula $B = \frac{S_x}{N + (x^2 - 1)}$. This benefit must then be divided by some measure of the cost. One indication of cost is the time spent in learning (T). So that

$$C/B = \sum \frac{\frac{S_x}{N + (x^2 - 1)}}{T}$$

Applying this formula to the data in Table 1, which is an expansion of the last figure in § 3.3.2 where ($sI = 12$), ($N = 10$) and ($T = 20$), we have ($s_1 = 6$), ($s_2 = 3$), ($s_3 = 2$), ($s_4 = 1$), so that:

$$\sum \frac{S_x}{N + (x^2 - 1)} = \frac{6}{10 + (1^2 - 1)} = \frac{3}{10 + (2^2 - 1)} = \frac{2}{10 + (3^2 - 1)} = \frac{1}{10 + (4^2 - 1)} = 6/10 + 3/13 + 2/18 + 1/25 = .882$$

But $T = 20$. Therefore $C/B = \sum \frac{\frac{S_x}{N + (x^2 - 1)}}{T} = .882/20 = .044$

If the polychronometer circuits are properly connected with some sort of computer, all the above calculations can be automated.

When making comparisons on the basis of such cost-benefit calculations, however, much precaution is necessary in order to assure and prove comparability. This is particularly true when comparing set of lessons in which a different foreign language is taught. In such cases the language differential must also be accounted for; some techniques for measuring these language differences have already been elaborated.

5. Toward More Efficient Language Teaching

Once language teaching behaviour and the corresponding language learning behaviour have been quantified it is then possible, as we have seen, to study the relationship between them. What remains is the practical and experimental problem of discovering either a cause-and-effect relationship or a significant statistical correlation between the highest benefit per cost, between the type of lesson profile and the type, order and proportion of the profile components. Between these and types of lesson plans, types of teachers and teaching styles and the materials used as expressed in quantified method profiles there are undoubtedly relationships which may be crucial. When such are established, it may be possible to study the ways in which modifications of a syllabus of language teaching may bring the language teaching process closer to the achievement of its stated objectives.

EPILOGUE

Is Language Teaching a Waste of Time?

The foregoing may lead to the following observations. Language lessons are different even when the same thing is being taught by the same method. The differences are due mainly to the behaviour of the teachers. The teaching method used does not, however, seem to make as much difference as the teaching behaviour. Most of the methods with speaking objectives produce little speech. The closeness of the lesson to the application of the theory has little relationship to the performance of the learner. Some of the lessons which we have analysed had been considered to be exemplary ones, conforming as they did to most of the tenets of this or that theory of language didactics - the sort of lessons one could use as models for the training of teachers. Yet they were not very much more productive than the rest. Can these observations be generalized to language teaching in most schools?

Satisfactory results have admittedly been achieved under certain circumstances in specialized civilian and military language schools with classes of six to eight highly motivated learners working six to eight hours a day under teams of competent and skilled teachers. Whether the same learners equally motivated and left alone with appropriate and abundant materials would achieve a better understanding of the language is a question which would merit serious investigation backed by some rigorous comparative research. For the implications are indeed momentous.

Such research is now in the realm of possibility. For we are now beginning to have archives of recorded language lessons, most on audiotape, but increasingly on videotape. If one listens to these lessons, most of them, in spite of great differences in approach and procedure, seem to have a common element. When the teacher is not talking, she or he is trying to extract utterances from the learners, the success rate seeming to vary from nil to poor - but rarely very good. In lessons where there is a lot of student talk, these utterances, most of them largely incorrect, are what all the other learners hear most of the time. In order to get more utterances from the learners during the brief time allotted to language teaching, teachers may divide the class into pairs or conversation groups. The more this is done the less the teacher is involved. The less teacher involvement the greater the speech productivity of the learners. Yet because of a lack of exposure to real models of the language spoken in real situations, learners are allowed to communicate in a sort of creolized 'interlanguage' which eventually becomes their version of the second language. At a later level a new course for teachers is divised whereby they are told to 'repair' the often irreparable damage. In either approach performance-oriented classroom learning creates stressful situations which become incompatible with the effective acquisition of a language.

When we look at the comprehension skills, the main function of the teacher seems to be to explain meaning, that is to act as a sort of involuntary glossary, often explaining the obvious and failing to satisfy the needs of many individuals, since not every learner grasps the meaning as quickly or as well as the others. It would seem that, time-wise, a specially prepared illustrated glossary to the materials being read or heard by the learner would be more efficient.

If the above assessment is correct, it leads to some disturbing hypotheses, the main one being: is language teaching a waste of time? Few of the millions of learners who are taught to speak a foreign language in school actually complete the course with an ability to speak the language. Although this is common knowledge, the failure is invariably attributed to poor teaching. This may indeed be a contributing factor. Yet the amount of language learning achieved in a public school system under the best teachers is not all that impressive. So what keeps the pot boiling?

Foreign languages have become the object of some well entrenched vested interests. Parents, politicians, publishers and educators have a belief that children should learn to speak a foreign language in school, preferably at an early age. So 'foreign language' is added to the list of school subjects which teachers are required to teach, even though many of them may have little knowledge of the language they are supposed to teach.

Some parents, remembering the poor results of their own language learning experience in school will see to it that their children are exposed to the language outside the school, often through periods of residence abroad, exchange programmes, periods of schooling in the other language and the like. The time spent learning the language in school is viewed as a useful preparation for these out-of-school learning activities. Most parents, however, wanting their children to learn a foreign language are satisfied if it is taught in school. The criterion of success is the result obtained in the examination. If Johnny is first in his language class he must be "good at languages". If Mildred is below average, she is not good at languages. Many of these results have little relationship with actual language speaking performance. They are simply based on achievment as defined by what has been taught, not necessarily on language competence as measured against the alleged speaking objective. Many if not most of the tests used to measure progress in courses with speaking objectives are actually comprehension tests.

Although most of this is evident, there is this persistent belief that a school population can learn to speak a foreign language by attending language lessons. If the population fails to do so after a few years, the blame is on the teachers. The teachers in turn blame the method. So there is a constant search for the right method of language teaching. For such things one tends to look to the sciences, the language sciences and the learning sciences. Teachers are therefore invited to listen to the scientific theories of specialists in linguistics, in psychology and in education who periodically will come and tell the school teachers of their discoveries and their findings which they believe might revolutionize language teaching. Teachers are told that "language is structure" so they must use a structural approach - pattern practice, speech drills and the like, where the form is stressed at the expense of the meaning. After of few years of this approach not producing the desired learning, another formula is proposed. "Language is communication" so one must use a communicative approach whereby learners are expected to communicate in a language they do not know things they cannot understand. One could almost predict the next wave. For these same approaches are bound to reappear periodically under different names, as they have ever since languages were first taught in school.

From time to time, the failure of school-based language learning has been documented in such reports as that of Coleman in the thirties and Dunkel in the forties. Coupled with years of direct observation of language lessons and of the performance of trained language teachers one can hazard a certain number of tentative conclusions.

Even the best language teachers, after years of frustration, become mediocre. They see how little they can achieve within the school system for the amount of effort expended. The system seems to crush even the most creative teachers; the best, it seems, can give no more than five years of peak performance. The first put-down is the textbook. "This is the book you will teach from for the coming year." Then there is the intimidation of the experts at conferences for language teachers. Then back to conformity to the programme, the time-consuming and often unnecessary testing, the paper work, and the large classes of unmotivated and often undisciplined learners. Under such circumstances learning to speak a language at the rate of a few hours a week would be almost miraculous. This speaking objective, because of its inefficient use of class time may actually deprive the learners of the possibility of acquiring knowledge in a language other than their own. It seems doubtful that taking a language as a school subject is the best way for people to learn to speak it. In language learning, the home seems to do better than the school; witness the level of competence achieved within literate bilingual families. So does the community. In multilingual neighborhoods children learn to speak another language by associating with those who speak it. In both home and street the other language is not taught; it is learned.

In sum, there seems to be no magic formula, no discipline that can provide one, no approach that can give all desired results. For language is not only structure, or communication, or speech, or whatever; it is all of these, and much more, the learning of which is complex, subtle, conscious and unconscious, achieved through both thought, memory, action and emotion, in more contexts than a classroom can provide. It is time to question whether teaching a language in school is really worth the time.

William F. Mackey

REFERENCES

ALBRECHT, Karl & Ron Kembe. *Service America: Doing Business in the New Economy*. Homewood, Ill.: Dow Jones-Irwin 1985.

AMIDON, Edmund & Ned A. Flanders. "The Effects of Direct and Indirect Influence on Dependent Prone Students Learning Geometry", in *Interaction Analysis: Theory, Research and Application*. Amidon and Hough, eds. Reading, Mass.: Addison-Wesley 1967, p. 210-216.

ANDERSON, H.H. & J.E. Brewer. "Studies of Teachers' Classroom Personalities. 1. Dominative and Socially Integrative Behaviour of Kindergarten Teachers", in *Applied Psychology Monograph* 6, 1945.

BIALYSTOK, E. "A Theoretical Model of Second Language Learning", in *Language Learning* 28 (1978) 1: 69-83.

BREEN, M.P. & C.N. Candlin. "The Essentials of a Communicative Curriculum in Language Teaching", in *Applied Linguistics* 2 (1981): 89-112.

BUITEN, R. & H. Lane. "A Self-Instructional Device for Conditioning Accurate Prosody", in *International Review of Applied Linguistics* 3 (1965): 205-219.

CHAPPLE, E.D. & G. Donald Jr. "An Evaluation of Department Store Salespeople by the Interaction Chronograph", in *The Journal of Marketing* 1947 (October): 173-185.

CHAPPLE, E.D. "The Interaction Chronograph", in *Personnel* 25 (1949): 295-307.

COGAN, M.L. "Theory and Design of a Study of Teacher-Pupil Interaction", in *The Harvard Educational Review* 26 (Fall 1956) 4: 315-342.

FLANDERS, Ned A. "Some Relationships Among Teacher Influence, Pupil Attitude and Achievements", in *Interaction Analysis: Theory, Research and Application*. Amidon and Hough, eds., p. 103-116.

-----. *Analyzing Teaching Behaviour*. Reading, Mass.: Addison-Wesley 1970.

FLIESS, R. (ed.). *The Psychoanalytic Reader*. New York: International University Press 1948.

GAGE, Nathaniel L., (ed.). *Handbook of Research on Teaching*. Chicago: Rand McNally 1967.

GOLDMAN-EISLER, F. "On the Variability of the Speed of Talking and its Relation to the Length of Utterances in Interviews", in *British Journal of Psychology* 45 (1954): 94-107.

HALBERSTAM, David. *The Reckoning: the Challenge to America's Greatness*. New York: Morrow 1986.

HAMMERLY, Hector. *An Integrated Theory of Language Teaching*. Burnaby: Second Language Publications 1985.

HEECHUNG, Veronica. *The Effects of Didactic and Language Behaviour on the Linguistic Performance of Eleven-Year-Olds in Five ESL Classrooms*. Quebec: Laval University M.A. Thesis, 1981.

HOCHSCHILD, Arlie Russel. *The Managed Heart: Commercialization of Human Feeling*. Berkeley: University of California Press 1983.

IRWIN, J.V. & O. Becklund. "Norms for Maximum Repetitive Rates for Certain Sounds Established with the Sylrater", in *Journal of Speech and Hearing Disorders* 18 (1953): 149-160.

JARVIS, Gilbert A. "A Behavioural Observation System for Classroom Foreign Language Skill Acquisition Activities", in *The Modern Language Journal* 52 (October 1968) 6: 335-341.

KASL, S.V. & G.F. Mahl. "A Simple Device for Obtaining Certain Verbal Activity Measures during Interviews", in *Journal of Abnormal and Social Psychology* 53 (1965): 425-433.

KRAMARAE, Chris (ed.). *Women and Men Speaking: Framework for Analysis.* Rowley: Newbury House, 1981.

KRASHEN, S. *Second Language Acquisition and Second Language Learning*. Oxford: Pergamon Press, 1981.

KRASNER, L. "Studies of the Conditioning of Verbal Behavior", in *Psychological Bulletin* 55 (1958): 148-170

LOUISY, Calliopa Perlette. *Polychronometry Compared with Flint in Analysing Foreign Language Teaching Behaviour*. Quebec: Laval University M.A. Thesis 1975.

LOVELESS, Glenn. *Four Issues in Language Didactics: An Investigation of the Effects of Specific Classroom Behaviours on the Learning of Third and Ninth Grade Students*. Quebec: Laval University Doctorate Thesis, 1983.

MACKEY, William Francis. *Language Teaching Analysis*. London: Longman 1965; Bloomington: Indiana University Press 1967, (French Version, *Principes de didactique analytique*. Paris: Didier 1972.) (Russian Version: Moscow 1973). (Japanese Version: Tokyo 1979).

-----. "La nuova tecnologia nella preparazione glottodidattica, in *Homo Loquens* 1 (1968a): 71-75 (trans. of W.F. Mackey, "The New Technology of Teacher Training, in *English Language Teaching* 23 (1968): 10-14).

-----. "Titocinio Didattico: modeli e moduli", in *Homo Loquens* 3 (1968b): 25-34 (trans. of W.F. Mackey, "Practice Teaching: Models and Modules", in *Les langues modernes* 3 (1968b):94-100).

-----. Sociolinguistics and the Synchronic Fallacy", in *Academiae Analecta* 45 (1983) 4: 79-90.

-----. "The Sociobiology of Ethnolinguistic Nucleation", in *Politics and the Life Sciences* 4 (1985): 10-14.

MARQUIS, Michel & Rosaire Saindon. *Introducing English - One*. Montreal: Holt, Rinehart and Winston 1962, Teacher's Text, 62 p., Student's Text, 31 p.

MATARAZZO, J.D., G. Saslow & R.G. Matarazzo. "The Interaction Chronograph, in *Journal of Psychology* 41 (1956): 347-367.

MEDLEY, D.M. & H.E. Mitzel. "A Technique for Measuring Classroom Behavior", in *Journal of Educational Psychology* 49 (April 1958) 2: 86-92.

-----. "Measuring Classroom Behavior by Systematic Observation", in *Handbook of Research on Teaching*, N.L. Gage, ed. Chicago: Rand McNally 1967, p. 247-328.

MELBIN, M. "An Interaction Recording Device for Participant Observers, in *Human Organization* 13 (1954) 2: 29-33.

MOSKOWITZ, Gertrude. *The Foreign Language Teacher Interacts.* Minnesota: Association for Productive Teaching, 1970.

-----. "The Effects of Training Foreign Language Teachers in Interaction Analysis", in *Foreign Language Annals* 1 (March 1968) 3: 218-235.

-----. "Interaction Analysis: A New Modern Language for Supervisors", in *Foreign Language Annals* 5 (December 1971) 2: 211-221.

POTTER, R.K., G.A. Kopp & H.C. Green. *Visible Speech.* New York: Van Nostrand 1947.

PUCKETT, R.C. "Making Supervision Objective", in *School Review* 26 (March 1928) 3: 209-212.

RIVERS, W. *Speaking in Many Tongues.* Rowley: Newbury House Publishers, 1972.

ROTHFARB, Sylvia H. "Teacher-Pupil Interaction in the FLES Class", in *Hispania* 53 (May 1970) 2: 256-260.

SAINSBURY, P. "A Method of Recording Spontaneous Movements by Time-Sampling Motion Pictures", in *Journal of Mental Science* 100 (1954): 742-748.

SHAIKEN, Harley. *Work Transformed: Automation and Labor in the Computer Age.* New York: Holt, Rinehart & Winston 1985.

STARKWEATHER, J.A. "A Speech Rate Meter for Vocal Behavioral Analysis, in *Journal of the Experimental Analysis of Behavior* 3 (1960): 111-114.

STERN, H.H. *Fundamental Concepts of Language Teaching.* New York: Oxford University Press 1983.

VERZEANO, M. & J.E. Finesinger. "An Automatic Analyzer for the Study of Speech in Interaction and in Free Association", in *Science* 110 (1949): 45-46.

WIENS, A.N., J.D. Matarazzo & G. Saslow. "The Interaction Recorder: an Electronic Punch Paper Tape Unit for Recording Speech Behavior During Interviews", in *Journal of Clinical Psychology* 21 (1965): 142-145.

WILSON, E.O. *Sociobiology: The New Synthesis.* Cambridge: Harvard University Press, 1975.

WRAGG, E.C. "Interaction Analysis in the Foreign Language Classroom", in *The Modern Language Journal* 54 (February 1970) 2: 116-120.

CODES AND SIGLA

AC	Action response
AL	All error types
AM	Active motivation
AN	Answer
CA	Class administration
CC	Correct after correction
CG	Correction of grammar
CH	Choral response
CO	Confirming response
CP	Correction of pronunciation
CR	Choral repetition
CS	Completion of learner's statement
DE	Description
DM	Dialogue modification
DO	Do (checking through doing)
DR	Dramatization / Duration
DS	Do and say
ER	Explaining rules
EX	Explanation
FP	Form presentation
FR	Form repetition / Frequency
FV	Form verification
GC	Guided conversation
GD	Giving dictation
GE	Grammatical error
GG	Guessing game
GV	Grammar and vocabulary mistake
IA	Interation (non-verbal)
IT	Iteration
KS	Look and say
KW	Look and write
LK	Linking (known to unknown)
LD	Listen and do
LR	Listen and repeat
LS	Listen and say
LW	Listen and write
MI	Miming
MM	Manipulating machine
MO	Manipulating object
MT	Mother tongue
MV	Movement (didactic)
OA	Other activities
OC	Other techniques
OD	Oral drills
OG	Oral games
OK	Completely correct utterance
OM	Oral modeling
OP	Other types of presentation
OR	Other types of repetition
OU	Other uses of language
OV	Other types of verification
PA	Presentation through actions
PE	Phonetic error / Pointing
PF	Presentation through films
PG	Pronunciation and grammar mistake
PO	Presentation through objects
PP	Pattern practice
PR	Prompting
PS	Presentation through speech
PT	Presentation through translation
PV	Pronunciation and vocabulary mistake
PX	Point to picture
QA	Question-and-asnwer drill
QL	Questioning learner
RC	Recite
RD	Reproduction drill
RE	Read aloud
RE	Re-explanation
RM	Re-modeling / Re-phrasing
RS	Read and say
RT	Repeat together
SA	Speech through actions
SI	Statement independently generated (by learner)
SH	Speech response (by learner)
SM	Speech and motion / Speech motivation
SP	Semantic presentation
SR	Semantic repetition
SS	Speech through speech
ST	Speech through translation
SV	Semantic verification
SX	Speech through pictures
TR	Translation
TT	Topical talk / Prepared speech
VE	Vocabulary error
WE	Write on board (learner)
WB	Writing on board (teacher)
WR	Write
WV	Written verification (write the answer)